

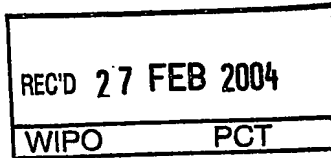


Best Available Copy



INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

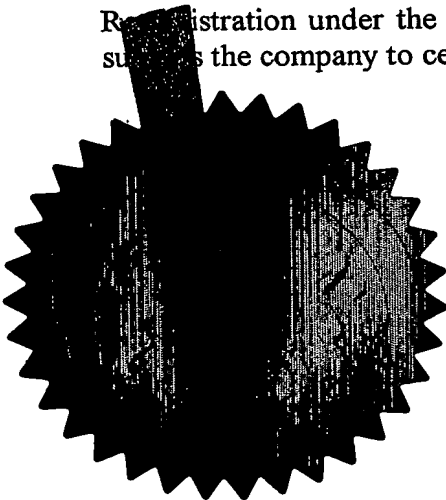


I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.



Signed

R. Mahoney

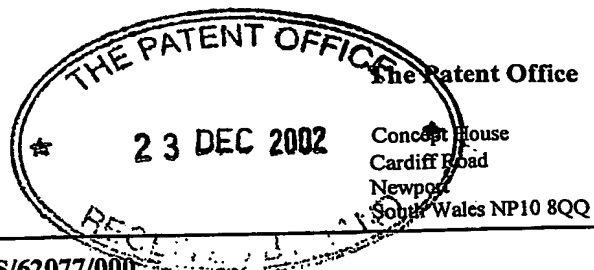
Dated 22 December 2003

**PRIORITY
DOCUMENT**
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

Best Available Copy

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)



1.	Your reference	SCB/CDM/MPS/62077/000		
2.	Patent application number (The Patent Office will fill in this part)	0230014.3		23 DEC 2002
3.	Full name, address and postcode of the or of each applicant (underline all surnames)	DevGen NV Technologiepark 9 B-9052 Ghent-Zwijnaarde Belgium		
	Patents ADP number (if you know it)			
	If the applicant is a corporate body, give the country/state of its incorporation	Belgium		
4.	Title of the invention	Kinase sequences useful for developing compounds for the prevention and/or treatment of metabolic diseases and nucleotide sequences encoding such kinase sequences		
5.	Name of your agent (if you have one)	BOULT WADE TENNANT		
	"Address for service" in the United Kingdom to which all correspondence should be sent (including the postcode)	VERULAM GARDENS 70 GRAY'S INN ROAD LONDON WC1X 8BT		
	Patents ADP number (if you know it)	42001 ✓		
6.	If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and (if you know it) the or each application number	Country	Priority application number (if you know it)	Date of filing (day/month/year)
7.	If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application	Number of earlier application	Date of filing (day / month / year)	
8.	Is a statement of inventorship and of right to grant of a patent required in support of this request? (Answer 'Yes' if: a) any applicant named in part 3 is not an inventor, or b) there is an inventor who is not named as an applicant, or c) any named applicant is a corporate body. See note (d))	Yes		

07454911009

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form

Description 89

Claim(s) 3

Abstract

Drawing(s) 2

10. If you are also filing any of the following, state how many against each item.

Priority documents

Translations of priority documents

Statement of inventorship and right to grant of a patent (Patents Form 7/77)

Request for preliminary examination and search (Patents Form 9/77) 1

Request for substantive examination (Patents Form 10/77)

Any other documents (Please specify)

11

I/We request the grant of a patent on the basis of this application.

Signature

Date

Baill Wade Tennant

23 December 2002

12. Name and daytime telephone number of person to contact in the United Kingdom

Claire Baldock
020 7430 7500

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

- If you need help to fill in this form or you have any questions, please contact the Patent Office on 01645 500505.
- Write your answers in capital letters using black ink or you may type them.
- If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.
- If you have answered 'Yes' Patents Form 7/77 will need to be filed.
- Once you have filled in the form you must remember to sign and date it.
- For details of the fee and ways to pay please contact the Patent Office.

DUPLICATE

Kinase sequences useful for developing compounds for the prevention and/or treatment of metabolic diseases and nucleotide sequences encoding such kinase sequences.

The present invention relates to nucleotide sequences that are useful in the pharmaceutical field.

In particular, the invention relates to nucleotide sequences that encode and/or may be used to express amino acid sequences that are useful in the identification and/or development of compounds with (potential) activity as pharmaceuticals, in particular of compounds for the prevention and/or treatment of metabolic diseases such as diabetes and obesity. These nucleotide sequences, which will be further described below, will also be referred to herein as "*nucleotide sequences of the invention*".

The invention also relates to the amino acid sequences - such as proteins and/or polypeptides - that are encoded by, and/or that may be obtained by suitable expression of, the nucleotide sequences of the invention. These amino acid sequences, which will be further described below, will also be referred to herein as "*amino acid sequences of the invention*".

The invention also relates to the use of the nucleotide sequences of the invention, preferably in the form of a suitable genetic construct as described below, in the transformation of host cells and/or host organisms, for example for the expression of the amino acid sequences of the invention. The invention also relates to host cells and/or host organisms that have been transformed with the nucleotide sequences of the invention and/or that can express the amino acid sequences of the invention.

The invention further relates to methods for the identification and/or development of compounds that can modulate the (biological) activity of the amino acid sequences of the invention, in which the abovementioned nucleotide sequences, amino acid sequences, genetic constructs, host cells and/or host organisms are used. Such methods which will usually be in the form of an assay or screen, as will also be further described below.

The invention also relates to the use of the nucleotide sequences, amino acid sequences, genetic constructs, host cells and/or host organisms of the invention in (methods for) identifying and/or developing compounds that can modulate the (biological) activity of the amino acid sequences of the invention.

Also, the invention also relates to compounds that can modulate the (biological) activity of the amino acid sequences of the invention, to compositions that contain such compounds, and to the use of such compounds in the preparation of such compositions.

In particular, the invention relates to such compositions that are in the form of pharmaceutical compositions, and more in particular in the form of pharmaceutical compositions for the prevention or treatment of metabolic diseases such as diabetes or obesity, and also to the use of compounds that can modulate the (biological) activity of the amino acid sequences of the invention in the preparation of such pharmaceutical compositions.

The invention further relates to the use of the nucleotide sequences, amino acid sequences, genetic constructs, host cells and/or host organisms of the invention in (methods for) identifying and/or developing compounds that can be used in the prevention or treatment of metabolic diseases.

Other aspects, embodiments, applications and advantages of the present invention will become clear from the further description below.

The present invention was established from the finding that the amino acid sequences of the invention are involved in metabolic processes (as further described below) and thus can be used as (potential) "target(s)" for *in vitro* and/or *in vivo* interaction with chemical compounds and other factors (with the term "target" having its usual meaning in the art, vide for example the definition given in WO 98/06737), and also from the finding that the nucleic acid sequence and amino acid sequences of the invention are involved in metabolic diseases. Consequently, compounds and/or factors that have been identified as interacting with the amino acid sequences of the invention (e.g. by the methods as described hereinbelow) may be useful as active agents in the pharmaceutical field, and in particular for the prevention and treatment of metabolic diseases. All this is supported by the following experimental data/observations:

- In an experimental model for fat handling, metabolism and storage using the model organism *C. elegans* (which model is further described in the Examples), downregulation of the *C. elegans* gene T17E9.1a by RNA interference strongly reduces the fat storage phenotype in said nematode. As further described below, T17E9.1a is the *C. elegans* ortholog of STE20 like kinase in mammals.

Some particularly preferred examples of nucleotide sequences of the invention are:

- the nucleotide sequence of SEQ ID NO:1 (T17E9.1a, also known as *kin-18*), which is a sequence derived from the nematode worm *C. elegans*; and
- the human orthologs of said *C.elegans* sequence, as may be identified by bioinformatic comparison of the *C.elegans* sequence with the human genome. Some preferred, but non-limiting orthologues are given in SEQ ID NOS: 3, 5, 7 and 9.

Generally herein, the use of the human nucleotide sequences of SEQ ID NOS: 3, 5, 7 and 9 and/or the use of nucleotide sequence derived therefrom will be preferred, in particular when the invention is used to develop compounds for pharmaceutical use.

In a broader sense, the term "*nucleotide sequence of the invention*" also comprises:

- parts and/or fragments of any of the nucleotide sequence of SEQ ID NO'S: 1, 3, 5, 7 and 9;
- (natural and/or synthetic) mutants, variants, alleles, analogs, orthologs (hereinbelow collectively referred to as "*mutants*") of any of the nucleotide sequence of SEQ ID NO'S: 1-5, as further described below.
- parts and/or fragments of such (natural or synthetic) mutants;
- nucleotide fusions of any of the nucleotide sequence of SEQ ID NO'S:1-5 (or a part or fragment thereof) with at least one further nucleotide sequence;
- nucleotide fusions of (natural or synthetic) mutants (or a part or fragment thereof) with at least one further nucleotide sequence;

in which such mutants, parts, fragments and/or fusions are preferably as further described below.

The invention also comprises different splice variants of the above nucleotide sequences.

Some particularly preferred examples of amino acid sequences of the invention are:

- the amino acid sequence of SEQ ID NO:2, which is a sequence derived from the nematode worm *C. elegans*; and

the human analogs of said *C.elegans* sequence, as may be identified by bioinformatic comparison of the *C.elegans* sequence with the human genome. Some preferred, but non-limiting analogs are given in SEQ ID NOS: 4, 6, 8 and 10.

In a broader sense, the term "*amino acid sequence of the invention*" also

comprises:

- parts and/or fragments of the amino acid sequence of SEQ ID NOS: 2, 4, 6, 8 and 10.
- (natural and/or synthetic) mutants, variants, alleles, analogs, orthologs (hereinbelow collectively referred to as "*analogs*") of the amino acid sequence of SEQ ID NOS: 2, 4, 6, 8 and 10;
- parts and/or fragments of such analogs;
- fusions of the amino acid sequence of SEQ ID NOS: 2, 4, 6, 8 and 10 (or a part or fragment thereof) with at least one further amino acid residue or sequence;
- fusions of the amino acid sequence of an analog (or a part or fragment thereof) with at least one further amino acid residue or sequence;

in which such mutants, parts, fragments and/or fusions are preferably as further described below.

The term "*amino acid sequence of the invention*" also comprises "immature" forms of the abovementioned amino acid sequences, such as a pre-, pro- or prepro-forms and/or fusions with suitable leader sequences. Also, the amino acid sequences of the invention may have been subjected to post-translational processing and/or be suitably glycosylated, depending upon the host cell or host organism used to express/produce said amino acid sequence; and/or may be otherwise modified (e.g. by chemical techniques known per se in the art).

Generally herein, the use of the human amino acid sequence of SEQ ID NOS: 4, 6, 8 and 10 and/or of amino acid sequences derived therefrom will be preferred, in particular when the invention is used to develop compounds for pharmaceutical use.

From the prior art, the following is known about the nucleotide sequences, amino acid sequences and compounds of the invention:

- (1) SEQ ID NOS: 1 and 2. These are the nucleotide sequence and amino acid sequence, respectively, of the *C. elegans* gene/protein T17E9.1a (also known as "*kin-18*")

Reference is made to the following entry from the WORMBASE database:

- for protein sequence:

<http://www.wormbase.org/db/seq/protein?name=WP%3ACE01405;class=Protein>

- for DNA sequence:

<http://www.wormbase.org/db/seq/sequence?name=T17E9.1a;class=Sequence>

as well as to Berman et al., Gene 279 (2001) 137-147.

- (2) SEQ ID NOS: 3 and 4. These are the nucleotide sequence and amino acid sequence, respectively, of the human gene/protein known as "*JNK/SAPK inhibitory kinase*" (also known as "*JIK*", "*DPK*" or "*KDS*", or STE20-like kinase).

Reference is made to the following entries from the GENBANK database:

- NM_016281(mRNA); and
- NP_057365.2 (translation).

See also:

- GENBANK entry AF135158; Zhang et al., Biochem. Biophys. Res. Commun. 274 (3), 872-879 (2000); MEDLINE 20384190, PUBMED 10924369; and
- GENBANK entry AF179867; Tassi et al., J. Biol. Chem. 274 (47), 33287-33295 (1999); MEDLINE 20026851; PUBMED 10559204

- (3) SEQ ID NOS: 5 and 6. These are the nucleotide sequence and amino acid sequence, respectively, of the human gene/protein known as "*prostate derived STE20-like kinase*" (also known as "*PSK*" or "*TAO2*").

Reference is made to the following entries from the GENBANK database:

- NM_016151.1 (mRNA)
- NP_057235.1 (translation)

See also:

- GENBANK entry AF061943; Moore et al, J. Biol. Chem. 275 (6), 4311-4322 (2000); MEDLINE 20127920; PUBMED 10660600;
- Chen et al., J. Biol. Chem., Vol. 274, No. 40, pp. 288803-28808 (1999).

- (4) SEQ ID NOS: 7 and 8. These are the nucleotide sequence and amino acid sequence, respectively, of the human gene/protein known as "*thousand and one amino acid protein kinase*" (also known as "*TAO1*").

Reference is made to the following entries from the GENBANK database:

- NM_004783 (mRNA)

- NP_004774.1 (translation)

See also:

- GENBANK entry AB020688; Nagase et al, DNA Res. 5 (6), 355-364 (1998);
MEDLINE 99156230; PUBMED 10048485

5 (5) SEQ ID NOS: 9 and 10. These are the nucleotide sequence and amino acid sequence, respectively, of the human gene/protein which with designation KIAA1361 (note that although this gene and protein are also known as "*thousand and one amino acid protein kinase*" or "*TAO1*", KIAA1361 is a different gene/protein from SEQ ID NOS 7 and 8).

10 Reference is made to the following entries from the EMBL and SpTrEMBL databases

- AY049015 (EMBL)
- Q9P2I6 (SPTREMBL)

See also: Nagase et al., DNA Res. 7:65 (2000) and EMBL entry NR AB037782 (<http://www.ebi.ac.uk/cgi-bin/emblfetch?AB037782>)

15 Also, although the inventors do not wish to be limited thereto, it is believed that SEQ ID NOS: 5/6 on the one hand, and SEQ ID NOS: 7 and 8 on the other hand, are transcripts from the same gene/locus.

Thus, in a first aspect, the invention relates to a nucleic acid, preferably in (essentially) isolated form, which nucleic acid encodes and/or can be used to express an
20 amino acid sequence of the invention (as defined herein), and in particular the amino acid sequence of SEQ ID NOS: 2, 4, 6, 8 or 10.

In another aspect, the invention relates to a nucleic acid, preferably in (essentially) isolated form, which nucleic acid comprises a nucleotide sequence of the invention, and in particular the nucleotide sequence of SEQ ID NOS: 1, 3, 5, 7 or 9.

25 In a yet another aspect, the invention relates to a nucleic acid, preferably in (essentially) isolated form, which nucleic acid essentially consists of a nucleotide sequence of the invention, and in particular a nucleotide sequence SEQ ID NO: 1, 3, 5, 7 or 9.

Collectively, these nucleic acids will also be referred to herein as "*nucleic acids of*
30 *the invention*". Also, where appropriate in the context of the further description of the

invention below, the terms "*nucleotide sequence of the invention*" and "*nucleic acid of the invention*" may be considered essentially equivalent and essentially interchangeable.

Also, for the purposes of the present invention, a nucleic acid or amino acid sequence is considered to "*(in) essentially isolated (form)*" – for example, from its native biological source - when it has been separated from at least one other component (and in particular macromolecule) with which it is usually associated, such as another nucleic acid, another protein/polypeptide or another (polymeric) biological component. In particular, a nucleic acid or amino acid sequence is considered "essentially isolated" when it has been purified at least 2-fold, in particular at least 10-fold, more in particular at least 100-fold, and up to 1000-fold or more.

The nucleic acids of the invention may also be in the form of a genetic construct, again as further described below. These constructs will also be referred to herein as "*genetic constructs of the invention*". In a preferred embodiment, such a construct will comprise:

- a) the nucleotide sequence of the invention; operably connected to:
 - b) one or more regulatory elements, such as a promoter and optionally a suitable terminator;
- and optionally also:
- c) one or more further elements of genetic constructs known per se;
- in which the terms "*regulatory element*", "*promoter*", "*terminator*", "*further elements*" and "*operably connected*" have the meanings indicated hereinbelow.

In another aspect, the invention relates to a protein or polypeptide, preferably in (essentially) isolated form, said protein or polypeptide comprising an amino acid sequence of the invention (as defined above), and in particular the amino acid sequence of SEQ ID NOS: 2, 4, 6, 8 or 10.

In a further aspect, the invention relates to a protein or polypeptide, preferably in (essentially) isolated form, said protein or polypeptide essentially consisting of an amino acid sequence of the invention (as defined above), and in particular of the amino acid sequence of SEQ ID NOS: 2, 4, 6, 8 or 10.

In a further aspect, the invention relates to methods for transforming a host cell and/or a host organism with a nucleotide sequence, with a nucleic acid and/or with a

genetic construct of the invention. The invention also relates to the use of a nucleotide sequence, of a nucleic acid and/or of a genetic construct of the invention transforming a host cell or a host organism.

In yet another aspect, the invention relates to a host cell or host organism that has been transformed and/or contains with a nucleotide sequence, with a nucleic acid and/or with a genetic construct of the invention. The invention also relates to a host cell and/or host organism that expresses, or (at least) is capable of expressing (e.g. under suitable conditions), an amino acid sequence of the invention. Collectively, such host cells/host organisms will also be referred to herein as "*host cells/host organisms of the invention*".

In yet another aspect, the invention relates to a methods for producing an amino acid sequence of the invention, in which a nucleotide sequence, nucleic acid, genetic construct, host cell or host organism of the invention is used. Such methods may for instance include expressing of a nucleotide sequence of the invention in a suitable host cell or host organism (e.g. upon suitable transformation), and/or maintaining and/or cultivating a host cell or host organism of the invention under suitable conditions, i.e. such that an amino acid sequence of the invention is expressed or obtained. Optionally, these methods may also comprise (one or more steps for) isolating the amino acid sequence thus expressed/produced. The invention also relates to the use of a nucleotide sequence, a nucleic acid, a genetic construct and/or a host cell/host organism of the invention in such a method.

In yet a further aspect, the invention relates to a method for identifying a compound that can modulate the (biological) activity of, and/or that can otherwise interact with, an amino acid sequence of the invention, which method is as further described below. The invention also relates to the use of a nucleotide sequence, a nucleic acid, a genetic construct, an amino acid sequence and/or a host cell/host organism of the invention in such a method.

In yet a further aspect, the invention relates to a method for identifying a compound that can be used in (the preparation of a pharmaceutical composition for) the prevention and/or treatment of metabolic diseases (as further defined below), which method is as further described below. The invention also relates to the use of a nucleotide

sequence, a nucleic acid, a genetic construct, an amino acid sequence and/or a host cell/host organism of the invention in such a method.

5 The invention also relates to compounds that can modulate the (biological activity of), and/or that can otherwise interact with, an amino acid sequence of the invention, either *in vitro* or preferably (also) *in vivo*, as further described below. The invention also relates to compositions that contain such compounds, and in particular to pharmaceutical compositions that contains such compounds.

10 The invention further relates to the use of compounds that can modulate the (biological activity of), and/or that can otherwise interact with, an amino acid sequence of the invention in the preparation of these compositions, and in particular to the use of such compounds in the preparation of a pharmaceutical composition for the prevention and/or treatment of metabolic diseases.

15 The invention also relates to compounds that can be used in the prevention and/or treatment of metabolic diseases (as further defined below), which compounds have or can be identified and/or developed using the method, nucleic acid sequence, amino acid sequence and/or host cell or host organism of the invention. The invention also relates to compositions that contain such compounds, and in particular to pharmaceutical compositions that contain said compounds.

20 The invention also relates to the use of such compounds in the preparation of a pharmaceutical composition, and in particular to the use of such compounds in the preparation of a pharmaceutical composition for the prevention or treatment of metabolic diseases.

25 Unless explicitly specified herein, all terms used in the present description have their usual meaning in the art, for which particular reference is made to the definitions given in WO98/06737 and EP 1 085 089.

The nucleotide sequences and amino acid sequences of the invention may generally further be characterized by the presence of a kinase domain (identified using SMART™-analysis, as follows:

30 - SEQ ID NO:2: from a.a. residue 30 to a.a. residue 289

-
- SEQ ID NO:4: from a.a. residue 24 to a.a. residue 277
 - SEQ ID NO:6: from a.a. residue 28 to a.a. residue 281
 - SEQ ID NO:8: from a.a. residue 28 to a.a. residue 281
 - SEQ ID NO:10: from a.a. residue 32 to a.a. residue 285

On the basis of the above and on the basis of the prior art referred to above, and although the invention is not specifically limited to any specific explanation or mechanism, it is assumed that the nucleotide sequences and/or amino acid sequences have (biological) activity as kinases.

In particular, and although the invention is again not limited to any specific explanation or hypothesis, it is assumed that the amino acid sequences of the invention may be involved in (the modulation of) the JNK/SAPK pathway (vide for example Zhang et al. and Tassi et al., both supra; as well as Yoneda et al., J. Biol. Chem., Vol. 276, No.17, p. 13935-13940 (2001)).

For JNKs and their role in diabetes/obesity, reference is inter alia made to Hirosumi et al., Nature. Vol. 20, November 2002, 333-336.

As is known in the art, biological activity of this kind can be measured using standard assay techniques for kinases, which are well known to the skilled person. Some preferred, but non-limiting examples include:

- the JNK/SAPK assay described by Zhang et al., supra, which is particularly suitable for the kinase of SEQ ID NO 4 or a similar kinase;
- the kinase assays described by Moore et al., supra, in particular the *in vitro* kinase assays described on page 4315;

The nucleotide sequence of SEQ ID NO: 1 was identified, and can be derived/isolated from/using the nematode *C.elegans*; in the manner as further described in Berman et al., or in any other suitable manner known per se.

The nucleotide sequences of SEQ ID NOS: 3, 5, 7 and 9 were identified, and can be derived/isolated from/using human cells; in the manner as further described in the prior art referred to above, or in any other suitable manner known per se.

Also, it is expected that - based upon the disclosure herein - the skilled person will be able to identify, derive and/or isolate natural "mutants" (as mentioned above) of the nucleotide sequence of SEQ ID NOS: 1, 3, 5, 7 and/or 9. For example, such mutants could be derived from (other individuals of) the same species (for example from an individual of a different strain or line, including but not limited to mutant strains or lines); and/or from (individuals of) other species (in which case these mutants will also be referred to herein as "orthologs"). Some examples of species from which such orthologs could be derived include, but are not limited to species of

- unicellular and/or micro-organisms such as bacteria, and yeast,
- invertebrate multicellular organisms as such as insects and nematodes (for example, agronomically harmful insect or nematode species);
- vertebrate multicellular organisms as such as fish, birds, reptiles, amphibians and mammals;

Preferably, a natural ortholog is derived from a mammal such as a mouse, rat, rabbit or dog.

Such natural mutants may be obtained in a manner essentially analogous to the methods described in the prior art referred to above, or alternatively by:

- construction of a DNA library from the species of interest in an appropriate expression vector system, followed by direct expression of the mutant sequence;
 - construction of a DNA library from the species of interest in an appropriate expression vector system, followed by screening of said library with a probe of the invention (as described below) and/or with a(nother) nucleotide sequence of the invention;
 - isolation of mRNA that encodes the mutant sequence from the species of interest, followed by cDNA synthesis using reverse transcriptase;
- and/or by any other suitable method(s) or technique(s) known per se, for which reference is for instance made to the standard handbooks, such as Sambrook et al, "Molecular Cloning: A Laboratory Manual" (2nd.ed.), Vols. 1-3, Cold Spring Harbor Laboratory

Press (1989) and F. Ausubel et al, eds., "Current protocols in molecular biology", Green Publishing and Wiley Interscience, New York (1987), as .

It is also expected that - based upon the disclosure herein - the skilled person will be able to provide and/or derive synthetic mutants (as defined hereinabove) of the nucleotide sequences of SEQ ID NOS: 1, 3, 5, 7 and/or 9.

Techniques for generating such synthetic sequences will be clear to the skilled person and may for instance include, but are not limited to, automated DNA synthesis; site-directed mutagenesis; combining two or more parts of one or more naturally occurring sequences, introduction of mutations that lead to the expression of a truncated expression product; introduction of one or more restriction sites (e.g. to create cassettes and/or regions that may easily be digested and/or ligated using suitable restriction enzymes), and/or the introduction of mutations by means of a PCR reaction using one or more "mismatched" primers, using for example a sequence of a naturally occurring GPCR as a template. These and other techniques will be clear to the skilled person, and reference is again made to the standard handbooks, such as Sambrook et al. and Ausubel et al., mentioned above.

Preferably, any mutants as described herein will encode amino acid sequences having one or more, and preferably all, of the structural characteristics/conserved features referred to above for the sequences of SEQ ID NO: 2, 4, 6, 8 and/or 10, and in particular will contain a kinase domain.

It is also possible in the invention to use a part or fragment of the nucleotide sequences of SEQ ID NOS 1, 3, 5, 7 or 9; or a part or fragment of a (natural or synthetic) mutant thereof. These may for instance be 5' and/or 3' truncated nucleotide sequences, or sequences with an introduced *in frame* startcodon or stopcodon. Also, two or more such parts or fragments of one or more nucleotide sequences of the invention may be suitably combined (e.g. ligated in frame) to provide a (further) nucleotide sequence of the invention.

Preferably, any such parts or fragments will be such that they comprise at least one continuous stretch of at least 15 nucleotides, preferably at least 30 nucleotides, more preferably at least 60 nucleotides, even more preferably more than 90 nucleotides, of one or more of the nucleotide sequences of SEQ ID NOS: 1, 3, 5, 7 and/or 9.

In particular, any mutants, parts or fragments as described herein may be such that they (at least) encode the active/catalytic site of the corresponding amino acid sequence of the invention and/or a binding domain of the corresponding amino acid sequence of the invention

5 Any mutants, parts and/or fragments as described herein are preferably (also) such that they are capable of hybridizing with one or more of the nucleotide sequences of SEQ ID NOS 1, 3, 5, 7 and/or 9, i.e. under conditions of "moderate stringency", and preferably under conditions of "high stringency". Such conditions will be clear to the skilled person, for example from the standard handbooks, such as Sambrook et al. and Ausubel et al.,
10 mentioned above, as well as in EP 0 967 284, EP 1 085 089 or WO 00/55318.

In particular, any mutants, parts and/or fragments as described herein may be such that they are capable of hybridizing with the nucleotide sequence of SEQ ID NO under the "stringent" hybridisation conditions described in WO 00/78972 and WO 98/49185, and/or under the hybridization conditions described in GB 2 357 768-A.

15 Also, any mutants, parts and/or fragments as described herein will preferably have a degree of "sequence identity", at the nucleotide level, with one or more of the nucleotide sequences of SEQ ID NOS: 1, 3, 5, 7 and/or 9 of at least 50%, preferably at least 60%, more preferably at least 70%, even more preferably at least 80%, and in particular more than 90%, and up to 95% or more.

20 For this purpose, the percentage of "sequence identity" between a given nucleotide sequence and the nucleotide sequence of SEQ ID NO:1 may be calculated by dividing [*the number of nucleotides in the given nucleotide sequence that are identical to the nucleotide at the corresponding position in the nucleotide sequence of the relevant SEQ ID NO*] by [*the total number of nucleotides in the given nucleotide sequence*] and
25 multiplying by [100%], in which each deletion, insertion, substitution or addition of a nucleotide - compared to the sequence of the relevant SEQ ID NO - is considered as a difference at a single nucleotide (position).

Alternatively, the degree of sequence identity may be calculated using a known computer algorithm for sequence alignment such as NCBI Blast v2.0, using standard
30 settings.

Some other techniques, computer algorithms and settings for determining the degree of sequence identity are for example described in EP 0 967 284, EP 1 085 089, WO 00/55318, WO 00/78972, WO 98/49185 and GB 2 357 768-A.

Also, in a preferred aspect, any mutants, parts and/or fragments as described herein will encode proteins/polypeptides having a biological activity that is essentially similar to the biological activity described above for the sequences of SEQ ID NOS: 2, 4, 6, 8 and/or 10, i.e. to a degree of at least 10%, preferably at least 50 % more preferably at least 75%, and up to 90%, as measured by the assay mentioned above. In particular, they will have activity as a kinase, as measured using a suitable assay of kinase activity, such as those referred to above, and preferably an activity which is at least 10%, preferably at least 50 % more preferably at least 75%, and up to 90% or more, of the kinase activity of any one of the sequences of SEQ ID NOS: 2, 4, 6, 8 and/or 10.

Preferably, any mutants, parts and/or fragments of the nucleotide sequence of the invention will (also) be such that they encode an amino acid sequence which has a degree of "sequence identity", at the amino acid level, with one or more of the amino acid sequence of SEQ ID NOS: 2, 4, 6, 8 and/or 10 of at least 50%, preferably at least 60%, more preferably at least 70%, even more preferably at least 80%, and in particular more than 90% and up to 95 % or more, in which the percentage of "sequence identity" is calculated as described below.

Preferably, a nucleotide sequence of the invention will (also) have a length (expressed as total number of nucleotides), which is at least 50%, preferably at least 60%, more preferably at least 70%, even more preferably at least 80%, and in particular more than 90% and up to 95 % or more of the length of one or more of the nucleotide sequence of SEQ ID NOS: 1, 3, 5, 7 and/or 9.

Generally, the nucleotide sequences of the invention, when in the form of a nucleic acid, may be DNA or RNA, and may be single stranded or double stranded. For example, the nucleotide sequences of the invention may be genomic DNA, cDNA or synthetic DNA (such as DNA with a codon usage that has been specifically adapted for expression in the intended host cell or host organism). Thus, the nucleotide sequences of the invention may contain intron sequences, and also generally comprises different splice variants.

It is also within the scope of the invention to use a fusion of a nucleotide sequence of the invention (as described above) with one or more further nucleotide sequence(s), including but not limited to one or more coding sequences, non-coding sequences and/or regulatory sequences. Preferably, in such fusions, the one or more further nucleotide sequences are operably connected (as described below) to the nucleotide sequence of the invention (for example so that, when the further nucleotide sequence is a coding sequence, the nucleotide fusion encodes a protein fusion as described below).

Another embodiment of the invention relates to a nucleic acid probe that is capable of hybridizing with a nucleotide sequence of the invention under conditions of moderate stringency, preferably under conditions of high stringency, and in particular under stringent conditions (all as described above). Such nucleotide probes may for instance be used for detecting and/or isolating a(nother) nucleotide sequence of the invention and/or as a primer for amplifying a nucleotide sequence of the invention; all using techniques known per se, for which reference is again made to the general handbooks such as Sambrook et al. and Ausubel et al. mentioned above.

Generally, such probes can be designed by the skilled person starting from a nucleotide sequence and/or amino acid sequence of the invention - and in particular one or more of the sequences of SEQ ID NOS 1, 3, 5, 7 and/or 9 - optionally using a suitable computer algorithm. Also, as will be clear to the skilled person, such probes may be degenerate probes.

In another embodiment, the invention relates to an antisense molecule against a nucleotide sequence of the invention.

Yet another embodiment relates to a double stranded RNA molecule directed against a nucleotide sequence of the invention (one strand of which will usually comprise at least part of a nucleotide sequence of the invention). The invention also relates to genetic constructs that can be used to provide such double stranded RNA molecules (e.g. by suitable expression in a host cell or host organism, or for example in a bacterial strain such as *E.coli*). For such constructs, reference is made to for example the International Applications PCT/IB01/1068 and WO 00/01846, both by applicant.

The amino acid sequence of SEQ ID NO: 2 was identified, and can be derived/isolated from/using the nematode *C.elegans*; in the manner as further described in Berman et al., or in any other suitable manner known per se.

The amino acid sequences of SEQ ID NOS: 4, 6, 8 and 10 were identified, and can be derived/isolated from/using human cells; in the manner as further described in the prior art referred to above, or in any other suitable manner known per se.

Generally, the amino acid sequences of SEQ ID NOS: 2, 4, 6, 8 and/or 10 may be isolated from the species mentioned above (i.e. *C. elegans* and *human*, respectively), using any technique(s) for protein isolation and/or purification known per se.

Alternatively, the amino acid sequences of SEQ ID NOS: 2, 4, 6, 8 and/or 10 may be obtained by suitable expression of a suitable nucleotide sequence - such as one of the nucleotide sequences of SEQ ID NOS: 1, 3, 5, 7 and/or 9, as applicable or a suitable mutant thereof - in an appropriate host cell or host organism, as further described below.

Also, it is expected that - based upon the disclosure herein - the skilled person will be able to identify, derive and/or isolate natural "analogs" (as mentioned above) of the amino acid sequences of SEQ ID NOS: 2, 4, 6, 8 and/or 10. Such mutants could be derived from (other individuals of) the same species (for example from an individual of a different strain or line, including but not limited to mutant strains or lines); and/or from (individuals of) other species. For example, such analogs could be derived from the insect species or other pest species mentioned above.

Such natural analogs may again be obtained by isolating them from their natural source using any technique(s) for protein isolation and/or purification known per se, or alternatively by suitable expression of a suitable nucleotide sequence of the invention - such as a natural mutant as described above - in an appropriate host cell or host organism, as further described below.

It is also expected that - based upon the disclosure herein - the skilled person will be able to provide and/or derive synthetic "analogs" (as mentioned above) of one or more of the amino sequences of SEQ ID NOS: 2, 4, 6, 8 and/or 10.

Generally, such synthetic analogs may be obtained by suitable expression of a suitable nucleotide sequence of the invention - such as a synthetic mutant as described above - in an appropriate host cell or host organism, as further described below.

Preferably, any analogs as described herein will have one or more, and preferably all, of the structural characteristics/conserved features referred to above for the sequences of SEQ ID NO: 2, 4, 6, 8 and/or 10, and in particular will contain a kinase domain.

It is also possible in the invention to use a part or fragment of one or more of the amino acid sequences of SEQ ID NOS 2, 4, 6, 8 and/or 10, or a part or fragment of a (natural or synthetic) analog thereof mutant thereof. This may for instance be N- and/or C- truncated amino acid sequence. Also, two or more parts or fragments of one or more amino acid sequences of the invention may be suitably combined to provide a (further) amino acid sequence of the invention.

Preferably, any such parts or fragments will be such that they comprise at least one continuous stretch of at least 5 amino acids, preferably at least 10 amino acids, more preferably at least 20 amino acids, even more preferably more than 30 amino acids, of one or more of the amino acid sequences of SEQ ID NO: 2, 4, 6, 8 and/or 10.

In particular, any parts or fragments as described herein are such that they (at least) comprise the active/catalytic site of the corresponding amino acid sequence of the invention and/or a binding domain of the corresponding amino acid sequence of the invention, and in particular a kinase domain. As will be clear to the skilled person, such parts or fragments may find particular use in assay- and screening techniques (as generally described below) and/or (when said part or fragment is provided in crystalline form) in X-ray crystallography.

Generally, such parts or fragments of the amino acid sequences of the invention may be obtained by suitable expression of a suitable nucleotide sequence of the invention - such as a suitable part or fragment as described hereinabove for the nucleotide sequences of the invention - in an appropriate host cell or host organism, as further described below.

In addition and/or as an alternative to the methodology above, amino acid sequences of the invention may also be provided by (chemically and/or enzymatically) modifying the side chain(s) of one or more amino acid residues of an amino acid sequence of SEQ ID NO: 2, 4, 6, 8 and/or 10 or a part, fragment, (natural and/or synthetic) mutant, variant, allele, analogs, orthologs thereof, for example by one or more of the side chain modifications as described in WO 01/02560 and/or by incorporating

(e.g. by insertion and/or substitution) one or more unnatural amino acid residues, again as described in WO 01/02560.

Preferably, any analogs, parts and/or fragments as described herein will be such that they have a degree of "sequence identity", at the amino acid level, with one or more of the amino acid sequences of SEQ ID NOS 2, 4, 6, 8 and/or 10 of at least 50%, preferably at least 60%, more preferably at least 70%, even more preferably at least 80%, and in particular more than 90% and up to 95 % or more.

For this purpose, the percentage of "sequence identity" between a given amino acid sequence and one of the amino acid sequences of SEQ ID NOS: 2, 4, 6, 8 and/or 10 may be calculated by dividing [*the number of amino acid residues in the given amino acid sequence that are identical to the amino acid residue at the corresponding position in the amino acid sequence of the relevant SEQ ID NO*] by [*the total number of amino acid residues in the given amino acid sequence*] and multiplying by [100%], in which each deletion, insertion, substitution or addition of an amino acid residue - compared to the sequence of the relevant SEQ ID NO - is considered as a difference at a single amino acid (position).

Alternatively, the degree of sequence identity may be calculated using a known computer algorithm, such as those mentioned above.

Also, such sequence identity at the amino acid level may take into account so-called "conservative amino acid substitutions", which are well known in the art, for example from GB-A-2 357 768, WO 98/49185; WO 00/46383 and WO 01/09300; and (preferred) types and/or combinations of such substitutions may be selected on the basis of the pertinent teachings from the references mentioned in WO 98/49185.

Also, preferably, any analogs, parts and/or fragments as described herein will have a biological activity that is essentially similar to the biological activity described above for the sequences of SEQ ID NOS: 2, 4, 6, 8 and/or 10, i.e. to a degree of at least 10%, preferably at least 50 % more preferably at least 75%, and up to 90%, as measured by the assay mentioned above. In particular, they will have activity as a kinase, as measured using a suitable assay of kinase activity, such as those referred to above, and preferably an activity which is at least 10%, preferably at least 50 % more preferably at

least 75%, and up to 90% or more, of the kinase activity of any one of the sequences of SEQ ID NOS: 2, 4, 6, 8 and/or 10.

Preferably, an amino acid sequence of the invention will (also) have a length (expressed as total number of amino acid residues), which is at least 50%, preferably at least 60%, more preferably at least 70%, even more preferably at least 80%, and in particular more than 90% and up to 95 % or more of the length of one or more of the amino acid sequence of SEQ ID NOS: 2, 4, 6, 8 and/or 10.

It is also within the scope of the invention to use a fusion of an amino acid sequence of the invention (as described above) with one or more further amino acid sequences, for example to provide a protein fusion. Generally, such fusions may be obtained by suitable expression of a suitable nucleotide sequence of the invention - such as a suitable fusion of a nucleotide sequence of the invention with one or more further coding sequences - in an appropriate host cell or host organism, as further described below.

One particular embodiment, such fusions may comprise an amino acid sequence of the invention fused with a reporter protein such as GFP, luciferase or another fluorescent protein moiety. As will be clear to the skilled person, such fusions may find particular use in expression analysis and similar methodologies.

In another embodiment, the fusion partner may be an amino acid sequence or residue that may be used in purification of the expressed amino acid sequence, for example using affinity techniques directed against said sequence or residue. Thereafter, said sequence or residue may be removed (e.g. by chemical or enzymatical cleavage) to provide the nucleotide sequence of the invention (for this purpose, the sequence or residue may optionally be linked to the amino acid sequence of the invention via a cleavable linker sequence). Some preferred, but non-limiting examples of such residues are multiple histidine residues and glutathione residues,

In one preferred, but non-limiting aspect, any such fusion will have a biological activity that is essentially similar to the biological activity described above for the sequences of SEQ ID NOS: 2, 4, 6, 8 and/or 10, i.e. to a degree of at least 10%, preferably at least 50 % more preferably at least 75%, and up to 90%, as measured by the assay mentioned above. In particular, they will have activity as a kinase, as measured

using a suitable assay of kinase activity, such as those referred to above, and preferably an activity which is at least 10%, preferably at least 50 % more preferably at least 75%, and up to 90% or more, of the kinase activity of any one of the sequences of SEQ ID NOS: 2, 4, 6, 8 and/or 10.

5 Genetic constructs of the invention will generally comprise at least one nucleotide sequence of the invention, optionally linked to one or more elements of genetic constructs known per se, as described below.

Such genetic constructs may be DNA or RNA, and are preferably double-stranded DNA. The constructs may also be in a form suitable for transformation of the intended
10 host cell or host organism, in a form suitable for integration into the genomic DNA of the intended host cell or in a form suitable independent replication, maintenance and/or inheritance in the intended host organism. For instance, the genetic construct may be in the form of a vector, such as for example a plasmid, cosmid, YAC, a viral vector or transposon. In particular, the vector may be an expression vector, i.e. a vector that can
15 provide for expression *in vitro* and/or *in vivo* (e.g. in a suitable host cell and/or host organism as described below).

As the one or more "further elements" referred to above, the genetic construct(s) of the invention may generally contain one or more suitable regulatory elements (such as a suitable promoter(s), enhancer(s), terminator(s), etc.), 3' - or 5' -UTR sequences, leader
20 sequences, selection markers, expression markers/reporter genes, and/or elements that may facilitate or increase (the efficiency of) transformation or integration. These and other suitable elements for such genetic constructs will be clear to the skilled person, and may for instance depend upon the type of construct used, the intended host cell or host organism; the manner in which the nucleotide sequences of the invention of interest are to
25 be expressed (e.g. via constitutive, transient or inducible expression); and/or the transformation technique to be used.

Preferably, in the genetic constructs of the invention, the one or more further elements are "*operably linked*" to the nucleotide sequence(s) of the invention and/or to each other, by which is generally meant that they are in a functional relationship with
30 each other. For instance, a promoter is considered "*operably linked*" to a coding sequence if said promoter is able to initiate or otherwise control/regulate the transcription and/or

the expression of a coding sequence (in which said coding sequence should be understood as being "*under the control of*" said promotor)

Generally, when two nucleotide sequences are operably linked, they will be in the same orientation and usually also in the same reading frame. They will usually also be essentially contiguous, although this may also not be required.

Preferably, the optional further elements of the genetic construct(s) used in the invention are such that they are capable of providing their intended biological function in the intended host cell or host organism.

For instance, a promoter, enhancer or terminator should be "*operable*" in the intended host cell or host organism, by which is meant that (for example) said promoter should be capable of initiating or otherwise controlling/regulating the transcription and/or the expression of a nucleotide sequence - e.g. a coding sequence - to which it is operably linked (as defined above).

Such a promoter may be a constitutive promoter or an inducible promoter, and may also be such that it (only) provides for expression in a specific stage of development of the host cell or host organism, and/or such that it (only) provides for expression in a specific cell, tissue, organ or part of a multicellular host organism.

Some particularly preferred promoters include, but are not limited to those present in the expression vectors referred to below.

A selection marker should be such that it allows - i.e. under appropriate selection conditions - host cells and/or host organisms that have been (successfully) transformed with the nucleotide sequence of the invention to be distinguished from host cells/organisms that have not been (successfully) transformed. Some preferred, but non-limiting examples of such markers are genes that provide resistance against antibiotics (such as kanamycine or ampicilline), genes that provide for temperature resistance, or genes that allow the host cell or host organism to be maintained in the absence of certain factors, compounds and/or (food) components in the medium that are essential for survival of the non-transformed cells or organisms.

A leader sequence should be such that - in the intended host cell or host organism - it allows for the desired post-translational modifications and/or such that it directs the transcribed mRNA to a desired part or organelle of a cell. A leader sequence may also

allow for secretion of the expression product from said cell. As such, the leader sequence may be any pro-, pre-, or prepro-sequence operable in the host cell or host organism.

An expression marker or reporter gene should be such that - in the host cell or host organism - it allows for detection of the expression of (a gene or nucleotide sequence present on) the genetic construct. An expression marker may optionally also allow for the localisation of the expressed product, e.g. in a specific part or organelle of a cell and/or in (a) specific cell(s), tissue(s), organ(s) or part(s) of a multicellular organism. Such reporter genes may also be expressed as a protein fusion with the amino acid sequence of the invention. Some preferred, but non-limiting examples include fluorescent proteins such as GFP.

For some (further) non-limiting examples of the promoters, selection markers, leader sequences, expression markers and further elements that may be present/used in the genetic constructs of the invention - such as terminators, transcriptional and/or translational enhancers and/or integration factors - reference is made to the general handbooks such as Sambrook et al. and Ausubel et al. mentioned above, to W.B. Wood et al., "*The nematode Caenorhabditis elegans*", Cold Spring Harbor Laboratory Press (1988) and D.L. Riddle et al., "*C. ELEGANS II*", Cold Spring Harbor Laboratory Press (1997), as well as to the examples that are given in WO 95/07463, WO 96/23810, WO 95/07463, WO 95/21191, WO 97/11094, WO 97/42320, WO 98/06737, WO 98/21355, US-A-6,207,410, US-A- 5,693,492 and EP 1 085 089. Other examples will be clear to the skilled person.

The genetic constructs of the invention may generally be provided by suitably linking the nucleotide sequence(s) of the invention to the one or more further elements described above, for example using the techniques described in the general handbooks such as Sambrook et al. and Ausubel et al., mentioned above.

Often, the genetic constructs of the invention will be obtained by inserting a nucleotide sequence of the invention in a suitable (expression) vector known per se. Some preferred, but non-limiting examples of suitable expression vectors include:

- vectors for expression in mammalian cells: pMAMneo (Clontech), pcDNA3 (Invitrogen), pMC1neo (Stratagene), pSG5 (Stratagene), EBO-pSV2-neo (ATCC 37593), pBPV-1 (8-2) (ATCC 37110), pdBPV-MMTneo (342-12) (ATCC 37224),

pRSVgpt (ATCC37199), pRSVneo (ATCC37198), pSV2-dhfr (ATCC 37146), pUCTag (ATCC 37460) and 1ZD35 (ATCC 37565);

- vectors for expression in bacterial cells: pET vectors (Novagen) and pQE vectors (Qiagen);
- 5 - vectors for expression in yeast or other fungal cells: pYES2 (Invitrogen) and Pichia expression vectors (Invitrogen);
- vectors for expression in insect cells: pBlueBacII (Invitrogen).

The nucleotide sequences and/or genetic constructs of the invention may be used to transform a host cell or host organism.

10 The host cell may be any suitable (fungal, prokaryotic or eukaryotic) cell or cell line, for example:

- a bacterial strain, including but not limited to strains of *E.coli*, *Bacillus*, *Streptomyces* and *Pseudomonas*;
- a fungal cell, including but not limited to cells from species of *Aspergillus* and *Trichoderma*;
- 15 - a yeast cell, including but not limited to cells from species of *Kluyveromyces* or *Saccharomyces*;
- an amphibian cell or cell line, such as *Xenopus* oocytes.

20 In one specific embodiment, which may particularly useful when the nucleotide sequences of the invention are (to be) used in the discovery and development of insecticidal compounds, the host cell may be an insect-derived cell or cell line, such as:

- cells/cell lines derived from *lepidoptera*, including but not limited to *Spodoptera* SF9 and Sf21 cells,
- cells/cell lines derived from *Drosophila*, such as Schneider and Kc cells; and/or
- 25 - cells/cell lines derived from a pest species of interest (as mentioned below), such as from *Heliothis virescens*.

In one preferred embodiment, the host cell is a mammalian cell or cell line, for example derived from the mammals referred to above.

30 In an even more preferred aspect, the host cell is a cell or cell line derived from a human, from the mammals including but not limited to CHO- and BHK-cells and human cells or cell lines such as HeLa and COS.

In one specific, but non-limiting embodiment, the cell or cell line may be human cell or cell line which is related to metabolic processes or metabolic disease and/or used as a cellular model for metabolic disease, including but not limited to liver cells or cell lines, adipocytes or muscle cells or cell lines such as HEPG2 cells, 3T3L1 adipocytes, CTC12 cells and L6 myotubes.

The host organism may be any suitable multicellular (vertebrate or invertebrate) organism, including but not limited to:

- a nematode, including but not limited to nematodes from the genus *Caenorhabditis*, such as *C.elegans*,
- an insect, including but not limited to species of *Drosophila* and/or a specific pest species of interest (such as those mentioned above);
- other well known model organisms, such as zebrafish;
- a mammal such as a rat or mouse;

Other suitable host cells or host organisms will be clear to the skilled person, for example from the handbooks and patent applications mentioned above.

It should be noted that when a nucleotide sequence of the invention is expressed in a multicellular organism, it may be expressed throughout the entire organism, or only in one or more specific cells, tissues, organs and/or parts thereof, for example by expression under the control of a promoter that is specific for said cell(s), tissue(s), organ(s) or part(s).

The nucleotide sequence may also be expressed during only a specific stage of development or life cycle of the host cell or host organism, again for example by expression under the control of a promoter that is specific for said stage of development or life cycle. Also, as already mentioned above, said expression may be constitutive, transient and/or inducible.

According to one specific embodiment, the expression of a nucleotide sequence of the invention in a host cell or host organism may be partly or totally reduced (i.e. knocked out), compared to the original (e.g. native) host cell or host organism. This may for instance be achieved in a transient manner using antisense and/or RNA-interference techniques well known in the art, or in a constitutive manner using random, site specific

and/or chemical mutagenesis of the nucleotide sequence of the invention, or any other suitable techniques for generating "knock-down" or "knock-out" animals.

Suitable transformation techniques will be clear to the skilled person and may depend on the intended host cell/host organism and the genetic construct to be used.

5 Some preferred, but non-limiting examples of suitable techniques include ballistic transformation, (micro-)injection, transfection (e.g. using suitable transposons), electroporation and lipofection. For these and other suitable techniques, reference is again made to the handbooks and patent applications mentioned above.

10 After transformation, a step for detecting and selecting those host cells or host organisms that have been successfully transformed with the nucleotide sequence/genetic construct of the invention may be performed. This may for instance be a selection step based on a selectable marker present in the genetic construct of the invention or a step involving the detection of the amino acid sequence of the invention, e.g. using specific antibodies.

15 The transformed host cell (which may be in the form of a stable cell line) or host organisms (which may be in the form of a stable mutant line or strain) form further aspects of the present invention.

20 Preferably, these host cells or host organisms are such that they express, or are (at least) capable of expressing (e.g. under suitable conditions), an amino acid sequence of the invention (and in case of a host organism: in at least one cell, part, tissue or organ thereof). The invention also includes further generations, progeny and/or offspring of the host cell or host organism of the invention, that may for instance be obtained by cell division or by sexual or asexual reproduction.

25 To produce/obtain expression of the amino acid sequences of the invention, the transformed host cell or transformed host organism may generally be kept, maintained and/or cultured under conditions such that the (desired) amino acid sequence of the invention is expressed/produced. Suitable conditions will be clear to the skilled person and will usually depend upon the host cell/host organism used, as well as on the regulatory elements that control the expression of the (relevant) nucleotide sequence of
30 the invention. Again, reference is made to the handbooks and patent applications mentioned above in the paragraphs on the genetic constructs of the invention.

Generally, suitable conditions may include the use of a suitable medium, the presence of a suitable source of food and/or suitable nutrients, the use of a suitable temperature, and optionally the presence of a suitable inducing factor or compound (e.g. when the nucleotide sequences of the invention are under the control of an inducible promoter); all of which may be selected by the skilled person. Again, under such conditions, the amino acid sequences of the invention may be expressed in a constitutive manner, in a transient manner, or only when suitably induced.

It will also be clear to the skilled person that the amino acid sequence of the invention may (first) be generated in an immature form (as mentioned above), which may then be subjected to post-translational modification, depending on the host cell/host organism used. Also, the amino acid sequence of the invention may be glycosylated, again depending on the host cell/host organism used.

The amino acid sequence of the invention may then be isolated from the host cell/host organism and/or from the medium in which said host cell or host organism was cultivated, using protein isolation and/or purification techniques known per se, such as (preparative) chromatography and/or electrophoresis techniques, differential precipitation techniques, affinity techniques (e.g. using a specific, cleavable amino acid sequence fused with the amino acid sequence of the invention) and/or preparative immunological techniques (i.e. using antibodies against the amino acid sequence to be isolated).

In one embodiment, the amino acid sequence thus obtained may also be used to generate antibodies specifically against said sequence or an antigenic part or epitope thereof.

Such antibodies, which form a further aspect of the invention, may be generated in a manner known per se, for example as described in GB-A-2 357 768, US-A- 5,693,492, WO 95/32734, WO 96/23882, WO 98/02456, WO 98/41633 and/or WO 98/49306, and/or as described in the prior art referred to above. Often, but not exclusively, such methods will involve as immunizing a immunocompetent host with the pertinent amino acid sequence of the invention or an immunogenic part thereof (such as a specific epitope), in amount(s) and according to a regimen such that antibodies against said amino acid sequence are raised, and then harvesting the antibodies thus generated, e.g. from blood or serum derived from said host.

For instance, polyclonal antibodies can be obtained by immunizing a suitable host such as a goat, rabbit, sheep, rat, pig or mouse with (an epitope of) an amino acid sequence of the invention, optionally with the use of an immunogenic carrier (such as bovine serum albumin or keyhole limpet hemocyanin) and/or an adjuvant such as Freund's, saponin, ISCOM's, aluminium hydroxide or a similar mineral gel, or keyhole limpet hemocyanin or a similar surface active substance. After a suitable immune response has been raised (usually within 1-7 days), the antibodies can be isolated from blood or serum taken from the immunized animal in a manner known per se, which optionally may involve a step of screening for an antibody with desired properties (i.e. specificity) using known immunoassay techniques, for which reference is again made to for instance WO 96/23882.

Monoclonal antibodies may for example be produced using continuous cell lines in culture, including hybridoma-based and similar techniques, again essentially as described in the above cited references. Accordingly, cells and cell lines that produce monoclonal antibodies against an amino acid sequence of the invention form a further aspect of the invention, as do methods for producing antibodies against amino acid sequences of the invention, which methods may generally involve cultivating such a cell and isolating the antibodies from the culture (medium), again using techniques known per se.

Also, Fab-fragments against the amino acid sequences of the invention (such as F(ab)₂, Fab' and Fab fragments) may be obtained by digestion of an antibody with pepsin or another protease, reducing disulfide-linkages and treatment with papain and a reducing agent, respectively. Fab-expression libraries may for instance be obtained by the method of Huse et al., 1989, Science 245:1275-1281.

In another embodiment, the amino acid sequence of the invention, or a host cell or host organism that expresses such an amino acid sequence, may also be used to identify and/or (further) develop compounds and/or other factors that can modulate the (biological) activity of, and/or that can otherwise interact with, the amino acid sequences of the invention, and such uses form further aspects of the invention. As will be clear to the skilled person, in this context, the amino acid sequence of the invention will serve as a target for interaction with such a compound or factor.

In this context, the terms "*modulate*", "*modulation*", "*modulator*" and "*target*" will have their usual meaning in the art, for which reference is *inter alia* made to the

definitions given in WO 98/06737. Generally, a modulator is a compound or factor that can enhance, inhibit/reduce or otherwise alter, influence or affect (collectively referred to as "*modulation*") a functional property of a biological activity or process (for example, the biological activity of an amino acid sequence of the invention).

5 In this context, the amino acid sequence of the invention may serve as a target for modulation *in vitro* (e.g. as part of an assay or screen) and/or for modulation *in vivo* (e.g. for modulation by a compound or factor that is known to modulate the target, which compound or factor may for example be used as an active compound for agrochemical, veterinary and/or pharmaceutical use).

10 For example, the amino acid sequences, host cells and/or host organisms of the invention may be used as part of an assay or screen that may be used to identify and/or develop modulators of the amino acid sequence of the invention, such as a primary screen (e.g. a screen used to identify modulators of the target from a set or library of test chemicals with unknown activity with respect to the target) and/or a secondary assay (e.g.
15 an assay used for validating hits from a primary screen and/or used in optimizing hit molecules, e.g. as part of hits-to-leads chemistry).

For instance, such an assay or screen may be configured as an *in vitro* assay or screen, which will generally involve binding of the compound or factor to be tested as a potential modulator for the target (hereinbelow also referred to as "test chemical") to the
20 target, upon which a signal generated by said binding is measured. Suitable techniques for such *in vitro* screening will be clear to the skilled person, and are for example described in Eldefrawi et al., (1987). FASEB J., Vol.1, pages 262-271 and Rauh et al., (1990), Trends in Pharmacol. Sci., vol.11, pages 325-329. For example, such an assay or screen may be configured as a binding assay or screen, in which the test chemical is used
25 to displace a detectable ligand from the target (e.g. a radioactive or fluorescent ligand), upon which the amount of ligand displaced from the target by the modulator is determined.

Assays aimed at identification of small molecule inhibitors of protein kinases are well known. There is a large number of technologies that can be classified in two main
30 approaches:

1. Detection based on an antibody specifically recognizing the phosphorylated product (peptide) of the kinase action.

These approaches rely on the specificity of antibody to differentiate between the peptide containing a phosphorylated amino acid vs. the non-phosphorylated peptide.

This recognition can then be used in the following readouts:

- fluorescence polarization (FP), where the detection is based on the change of the overall size upon antibody binding to fluorescent labeled peptide.
- fluorescence intensity (FI), where the antibody binding causes change in the intensity of the emission from a fluorescently labeled peptide.
- time-resolve fluorescence (where both the antibody and the peptide are labeled with different fluorescent labels and the binding leads to energy transfer between these two labels. This in turn leads to shortening of the donor label fluorescence lifetime. These assays have at least two different embodiments: biochemical or cell-based, via enzyme complementation (the latter is marketed as HitHunter (TM), by Discoverex).

2. Detection based on other reagents specifically recognizing phosphate groups.

Recently, at least three new techniques have been developed that do not rely on the use of specific antibody.

- Trivalent metal ion containing beads. These bind specifically to phosphate groups. The detection is then performed by change in fluorescence polarization, the same as above upon antibody binding. (IMAP(TM), Molecular Devices))
- based on chemical modification of the phosphate group and a specific label in the substrate to yield a change in fluorescence intensity. (IQ(TM), Pierce)
- based on chemical modification of the phosphate group to yield a biotin derivative. The detection is performed via FP upon avidin or streptavidin binding. (Caliper)"

Such an assay or screen may also be configured as a cell-based assay or screen, in which a host cell of the invention is contacted with/exposed to a test chemical, upon which at least one biological response by the host cell is measured.

Suitable cells or cell lines for such cell based assays include those mentioned above. In one preferred, but non-limiting embodiment, the cell or cell line may be a

mammalian, and in particular human, cell or cell line which is related to metabolic processes or metabolic disease and/or used as a cellular model for metabolic disease, including but not limited to liver cells or cell lines, adipocytes or muscle cells or cell lines such as HEPG2 cells, 3T3L1 adipocytes, CTC12 cells and L6 myotubes.

5 Also, such an assay or screen may also be configured as an whole animal screen, in which a host organism of the invention is contacted with/exposed to a test chemical, upon which at least one biological response (such as a phenotypical, behavioural and/or physiological change, including but not limited to paralysis or death) by the host organism is measured. Such screens may be carried out in any model organism known per se,
10 including but not limited to yeast, *Drosophila*, zebrafish or *C. elegans*.

 Thus, generally, the assays and screens described above will comprise at least one step in which the test chemical is contacted with the target (and/or with a host cell or host organism that expresses the target), and in particular in such a way that a signal is generated that is representative for the modulation of the target by the test chemical. In a
15 further step, said signal may then be detected.

 Accordingly, in one aspect, the invention relates to a method for generating a signal that is representative for the interaction of an amino acid sequence of the invention with a test chemical, said method at least comprising the steps of:

- 20 a) contacting the amino acid sequence of the invention, or a host cell or host organism containing/expressing an amino acid sequence of the invention, with said test chemical, in such a way that a signal may be generated that is representative for the interaction between said test chemical and said amino acid sequence; and optionally
 b) detecting the signal that may thus be generated.

 In another aspect, the invention relates to a method for identifying modulators of
25 an amino acid sequence of the invention (e.g. from a set or library of test chemicals), said method at least comprising the steps of:

- a) contacting the amino acid sequence of the invention, or a host cell or host organism containing/expressing an amino acid sequence of the invention, with a test chemical, in such a way that a signal may be generated that is representative for the interaction
30 between said test chemical and said amino acid sequence; and optionally

- b) detecting the signal that may thus be generated, said signal identifying a modulator of said amino acid sequence.

Compounds that may be tested using the methods of the invention are generally described below.

5 The assays and screens of the invention may be carried out at medium throughput to high throughput, for example in an automated fashion using suitable robotics. In particular, in this embodiment, the method of the invention may be carried out by contacting the target with the test compound in a well of a multi-well plate, such as a standard 24, 96, 384, 1536 or 3456 well plate.

10 Usually, in a screen or assay of the invention, for each measurement, the target or host cell/host organism will be contacted with only a single test compound. However, it is also within the scope of the invention to contact the target with two or more test compounds - either simultaneously or sequentially - for example to determine whether said combination provides a synergistic effect.

15 Once a test chemical has been identified as a modulator for an amino acid sequence of the invention (e.g. by means of a screen or assay as described hereinabove), it may be used per se as a modulator of the amino relevant amino acid sequence of the invention (e.g. as an active substance for pharmaceutical use), or it may optionally be further optimized for final use, e.g. to improve properties such as solubility, ADME-TOX
20 and other desired properties. It will be clear to the skilled person that the nucleotide sequences, amino acid sequences, host cells/host organisms and/or methods of the invention may find further use in such optimization methodology, for example as (part of) secondary assays.

25 The invention is not particularly limited to any specific manner or mechanism in/via which the modulator (e.g. the test chemical, compound and/or factor) modulates, or interacts with, the target (*in vivo* and/or *in vitro*). For example, the modulator may be an agonist, an antagonist, an inverse agonist, a partial agonist, a competitive inhibitor, a non-competitive inhibitor, a cofactor, an allosteric inhibitor or other allosteric factor for the target, and/or may be a compound or factor that enhances or reduces binding of target to
30 another biological component associated with its (biological) activity, such as another protein or polypeptide, a receptor, or a part of organelle of a cell. As such, the modulator

may bind with the target (at the active site, at an allosteric site, at a binding domain and/or at another site on the target, e.g. covalently or via hydrogen bonding), block the active site of the target (in a reversible, irreversible or competitive manner), block a binding domain of the target (in a reversible, irreversible or competitive manner), and/or influence or change the conformation of the target.

As such, the test chemical/modulator may for instance be:

- an analog of a known substrate of the target;
- an oligopeptide, e.g. comprising between 2 and 20, preferably between 3 and 15 amino acid residues;
- an antisense or double stranded RNA molecule;
- a protein, polypeptide;
- a cofactor or an analog of a cofactor.

Preferably, the compound is an inhibitor of the target, although the invention in its broadest sense is not limited thereto.

The test chemical/modulator may also be a reference compound or factor, which may be a compound that is known to modulate or otherwise interact with the target (e.g. a known substrate or inhibitor for the target) or a compound or factor that is generally known compound that is known to modulate or otherwise interact with other members from the general class to which the target belongs (e.g. a known substrate or inhibitor of said class).

Preferably, however, the compound(s) will be "small molecules", by which is generally meant herein a molecular entity with a molecular weight of less than 1500, preferably less than 1000. This may for example be an organic, inorganic or organometallic molecule, which may also be in the form or a suitable salt, such as a water-soluble salt; and may also be a complex, chelate and/or a similar molecular entities, as long as its (overall) molecular weight is within the range indicated above.

In a preferred embodiment, such a "small molecule" has been designed according, and/or meets the criteria of, at least one, preferably at least any two, more preferably at least any three, and up to all of the so-called Lipinski rules for drug likeness prediction (vide Lipinski et al., Advanced Drug Delivery Reviews 23 (1997), pages 3-25). As is known in the art, small molecules which meet these criteria are particularly suited (as

starting points) for the design and/or development of pharmaceuticals for human use, and may for instance be used as starting points for hits-to-leads chemistry, and/or as starting points for lead development (in which the methods of the invention may also be applied).

Also, for these purposes, the design of such small molecules (as well as the design of libraries consisting of such small molecules) will preferably also take into account the presence of pharmacophore points, for example according to the methods described by I. Muegge et al., J. Med. Chem. 44, 12 (2001), pages 1-6 and the documents cited herein.

The term "small peptide" generally covers (oligo)peptides that contain a total of between 2 and 35, such as for example between 3 and 25, amino acids (e.g. in one or more connected chains, and preferably a single chain). It will be clear that some of these small peptides will also be included in the term small molecule as used herein, depending on their molecular weight.

In one preferred, but non-limiting embodiment, the invention is used to screen a set or library of (related or otherwise unrelated) small molecules, for example a standard "robustness set", a primary screening library (e.g. of otherwise unrelated compounds), a combinatorial library, a series of closely related chemical analogos. Such sets or libraries will be clear to the skilled person, and may for instance include, but are not limited to, such commercially available chemical libraries such as the various libraries available from Tocris Cookson, Bristol, UK.

In yet a further aspect, the invention relates to a method for identifying a compound that can be used in (the preparation of a pharmaceutical composition for) the prevention and/or treatment of metabolic diseases (e.g. from a set or library of test chemicals), said method at least comprising the steps of:

- a) contacting an amino acid sequence of the invention, and/or a host cell or host organism containing/expressing an amino acid sequence of the invention, with a test chemical, in such a way that a signal may be generated that is representative for the interaction between said test chemical and said amino acid sequence of the invention; and optionally
- b) detecting the signal that may thus be generated, said signal identifying a modulator of said amino acid sequence.

The modulators thus identified can be used in (the preparation of a pharmaceutical composition for) the prevention and/or treatment of metabolic diseases, and/or can be used to develop other compounds that can be used in (the preparation of a pharmaceutical composition for) the prevention and/or treatment of metabolic diseases, i.e. as already outlined above.

The invention also relates to the use of an amino acid sequence and/or a host cell/host organism of the invention in such a method.

Also, as already mentioned above, the use of the human nucleotide sequences of SEQ ID NOS: 3, 5, 7 and/or 9 and/or of sequences derived therefrom (such as mutants, parts, fragments and/or fusions thereof as described hereinabove), of the human amino acid sequences of SEQ ID NOS: 4, 6, 8 and/or 10 and/or of sequences derived therefrom (such as analogs, parts, fragments, and/or fusions thereof as described hereinabove), and of host cells/host organisms containing/expressing these, are usually preferred, in particular when the invention is used to develop compounds for pharmaceutical use.

As already mentioned above, the compounds and/or factors that have been identified and/or developed as modulators of the amino acid sequences of the invention (and/or precursors for such compounds) may be useful as active substances in the pharmaceutical field, for example in the preparation of pharmaceutical compositions, and both such modulators as well as (pharmaceutical) compositions containing them further aspects of the invention.

In particular, the compounds and composition of the invention may be used in (the preparation of pharmaceutical compositions for) the prevention (e.g. prophylaxis) and/or treatment of metabolic diseases (which for the purposes herein in its broadest sense also includes preventing, treating and/or alleviating the symptoms and/or complications of such metabolic diseases).

In particular, such compounds and composition may be used in (the preparation of pharmaceutical compositions for) the prevention (e.g. prophylaxis) and/or treatment of metabolic diseases (which for the purposes herein in its broadest sense also includes preventing, treating and/or alleviating the symptoms and/or complications of such metabolic diseases).

In particular, the compounds and compositions of the invention may be used for preventing and/or treating:

- hyperglycemic conditions and/or other conditions and/or diseases that are (primarily) associated with (the response or sensitivity to) insulin, including but not limited to all forms of diabetes and disorders resulting from insulin resistance, such as Type I and Type II diabetes, as well as severe insulin resistance, hyperinsulinemia, and hyperlipidemia, e.g., obese subjects, and insulin-resistant diabetes, such as Mendenhall's Syndrome, Werner Syndrome, leprechaunism, lipotrophic diabetes, and other lipotrophies;
- conditions caused or usually associated with hyperglycemic conditions and/or obesity, such as hypertension, osteoporosis and/or lipodystrophy.
- so-called "metabolic syndrome" (also known as "Syndrome X") which is a condition where several of the following conditions coexist: hypertension; insulin resistance; diabetes; dyslipidemia; and/or obesity.

In particular, the compounds and compositions of the invention may be used for preventing and/or treating diabetes, especially Type I and Type II diabetes. "Diabetes" itself refers to a progressive disease of carbohydrate metabolism involving inadequate production or utilization of insulin and is characterized by hyperglycemia and glycosuria.

It should also be noted that, for the treatment of the metabolic disease in humans, the compound used will usually and preferably be an inhibitor of an amino acid sequence of the invention, although the invention in its broadest sense is not limited thereto.

In one specific, but non-limiting, embodiment of the invention, a compound is considered an inhibitor of one of the amino acid sequences of the invention if, in a relevant assay such as the kinase activity assays referred to above (or a suitable modification thereof, for example using partially or fully purified protein), said compound reduces the activity of said amino acid sequence, i.e. by at least 1%, preferably at least 10%, such as by 20% or more, compared to the activity without the presence of said compound.

In an even more specific, but non-limiting, embodiment of the invention, a compound is considered an inhibitor of one of the amino acid sequences of the invention if, in a relevant assay, such as a binding assay, said compound has an IC₅₀ value of less

than 1000 μm , preferably at than 500 μm , more preferably less than 250 μm , even more preferably less than 100 μm , for example 50 μm or less, such as about 10 μm or less.

Again, preferably, in the invention compounds are used that are modulators, and in particular inhibitors, of the human amino acid sequences of SEQ ID NO: 4, 6, 8 and/or 10, and/or of amino acid sequences derived therefrom, such as analogs, mutants, parts, fragments and/or fusions as described above.

For pharmaceutical use, the compounds of the invention may be used as a free acid or base, and/or in the form of a pharmaceutically acceptable acid-addition and/or base-addition salt (e.g. obtained with non-toxic organic or inorganic acid or base), in the form of a hydrate, solvate and/or complex, and/or in the form or a pre-drug, such as an ester. Such salts, hydrates, solvates, etc. and the preparation thereof will be clear to the skilled person; reference is for instance made to the salts, hydrates, solvates, etc. described in US-A-6,372,778, US-A-6,369,086 and US-6,369,067

Generally, for pharmaceutical use, the compounds of the inventions may be formulated as a pharmaceutical preparation comprising at least one compound of the invention and at least one pharmaceutically acceptable carrier, diluent or excipient and/or adjuvant, and optionally one or more further pharmaceutically active compounds. By means of non-limiting examples, such a formulation may be in a form suitable for oral administration, for parenteral administration (such as by intravenous, intramuscular or subcutaneous injection or intravenous infusion), for topical administration, for administration by inhalation, by a skin patch, by an implant, by a suppository, etc.. Such suitable administration forms - which may be solid, semi-solid or liquid, depending on the manner of administration - as well as methods and carriers for use in the preparation thereof, will be clear to the skilled person; reference is again made to for instance US-A-6,372,778, US-A-3,696, 086 and US-6,369,067.

The pharmaceutical preparations of the invention are preferably in a unit dosage form, and may be suitably packaged, for example in a box, blister, vial, bottle, sachet, ampoule or in any other suitable holder or container (which may be properly labeled); optionally with one or more leaflets containing product information and/or instructions for use. Generally, such unit dosages will contain between 1 and 500 mg of the at least

one compound of the invention, e.g. about 10, 25, 50, 100, 200, 500 or 1000 mg per unit dosage.

For pharmaceutical use, at least one compound of the invention will generally be administered in an amount of between 0.01 to 150 mg/kg body weight per day of the patient, divided over one or more daily doses. The amount(s) to be administered and the further treatment regimen may be determined by the treating clinician, depending on factors such as the age, gender and general condition of the patient and the nature and severity of the disease/symptoms to be treated.

Thus, in a further aspect, the invention relates to a composition, and in particular a composition for pharmaceutical use, that contains at least one compound of the invention (i.e. a compound that has been identified, discovered and/or developed using a nematode or method as described herein) and at least one suitable carrier (i.e. a carrier suitable for pharmaceutical use). The invention also relates to the use of a compound of the invention in the preparation of such a composition.

Preferably, the compounds and compositions of the invention are administered orally and/or in a form intended and/or suitable for oral administration.

It is also envisaged that the above compounds and compositions may be of value in the veterinary field, which for the purposes herein not only includes the prevention and/or treatment of diseases in animals, but also - for economically important animals such as cattle, pigs, sheep, chicken, fish, etc. - enhancing the growth and/or weight of the animal and/or the amount and/or the quality of the meat or other products obtained from the animal. Thus, in a further aspect, the invention relates to a composition for veterinary use that contains at least one compound of the invention (i.e. a compound that has been identified, discovered and/or developed using a nematode or method as described herein) and at least one suitable carrier (i.e. a carrier suitable for veterinary use). The invention also relates to the use of a compound of the invention in the preparation of such a composition.

In the agrochemical field, the invention may be used to identify compounds suitable for use in pesticides, insecticides, nematocides and/or other biocides or plant protection agents. For example, the compounds invention may be used to control the species listed in US-A-6,372,774. For this purpose, the compounds of the invention (or a

suitable salt, hydrate or ester thereof) may be suitably formulated with one or more agrochemically acceptable carriers, to provide a formulation suitable for agrochemical use, as will be clear to the skilled person (reference is for example made to the formulations and uses described in US-A-6,372,774).

Thus, in a further aspect, the invention relates to a composition for agrochemical use that contains at least one compound of the invention (i.e. a compound that has been identified, discovered and/or developed using a nematode or method as described herein) and at least one suitable carrier (i.e. a carrier suitable for agrochemical use). The invention also relates to the use of a compound of the invention in the preparation of such a composition.

The invention will now be further illustrated by means of the following non-limiting Experimental Part.

In the Figures:

- Figure 1 schematically shows vector pGN49A (see also WO 00/01846 and British patent application 0012233, both by Applicant);
- Figures 2A and 2B are a photographs (enhanced using the Scion Image (Scion Corp) software package) showing reduced fat-absorption phenotype in *C. elegans* upon Nile Red Staining: Figure 2A = reduced fat storage (T17E9.1a- invention); Figure 2B = reference (vector gGN29 without T17E9.1a).

Experimental part:

In the Experimental Part below, unless indicated otherwise, all steps for handling and cultivating *C. elegans* were performed using standard techniques and procedures, for which reference is made to the standard *C. elegans* handbooks, such as W.B. Wood et al., "*The nematode Caenorhabditis elegans*", Cold Spring Harbor Laboratory Press (1988); D.L. Riddle et al., "*C. ELEGANS II*", Cold Spring Harbor Laboratory Press (1997); "*Caenorhabditis elegans, Modern Biological analysis of an organism*": ed. by H. Epstein and D. Shakes, Methods in Cell Biology, Vol 48, 1995; and "*C. elegans, a practical approach*", ed. by I.A. Hope, Oxford University Press Inc. New York, USA, 1999.

Downregulation of the gene(s) of interest in *C. elegans* was achieved by RNAi feeding techniques using an *E.coli* strain capable of expressing a dsRNA corresponding

to the gene(s) of interest, as generally described in - *inter alia* - the International application WO 00/01846 by applicant and the handbooks referred to above.

Also, unless indicated otherwise, all cloning and other molecular biology steps were performed using standard techniques and protocols, i.e. as provided by the manufacturers of the reagents/kits used and/or as described in the standard handbooks, such as Sambrook et al, "Molecular Cloning: A Laboratory Manual" (2nd.ed.), Vols. 1-3, Cold Spring Harbor Laboratory Press (1989) and F. Ausubel et al, eds., "Current protocols in molecular biology", Green Publishing and Wiley Interscience, New York (1987).

Fat accumulation in *C. elegans daf-2 (e1370)* was determined visually under a microscope upon staining with Nile-red, using an adaptation of the general methodology described by Ogg et al., Nature, Vol. 389, 994 (1997). For the general methodology, reference is also made to Thaden et al., 1999 International Worm Meeting abstract 837; Ashrafi and Ruvkun, 2000 East Coast Worm Meeting abstract 67; Ashrafi, Chang and Ruvkun, 2001 International Worm Meeting abstract 325; and Rottiers and Antebi, 2001 International Worm Meeting abstract 620 (all abstracts available from Worm Literature Index at <http://elegans.swmed.edu/wli/>).

Example 1: Preparation *E.coli* RNA feeding strain for expression of T17E9.1a double stranded RNA.

A vector for expression of dsRNA for downregulation of *C. elegans* gene T17E9.1a (*kin-18*) was prepared as follows.

The DNA fragment of SEQ ID NO:11, which corresponds to 1576 nucleotides (26 % compared to the genomic sequence) of the *C. elegans* T17E9.1a gene, was obtained by PCR from genomic *C. elegans* DNA, using the following primers:

- forward primer : CGAAAACCAGCAGAAGAGCG [SEQ ID NO: 12]
- reverse primer : TCGAGGCATGTTTCAGACTTCG [SEQ ID NO: 13]

This fragment was inserted in the *SrfI*-site of expression vector pGN49a pGN49A (Figure 1, see also WO 00/01846 and British patent application 0012233, both by

Applicant). This vector contains two T7 promoters flanking the *SrfI*-site, allowing transcription of a nucleotide sequence inserted into said *SrfI*-site into double stranded RNA, upon binding of a T7 polymerase to said promoter (vide WO 00/01846).

The resulting vector, designated pGN49A-T17E9.1a, was transformed overnight into *E. coli* strain AB 309-105 (see EP-A-1 093 526 by applicant, page 17.).

To normalize the culture, 250 µl of the overnight culture (1 ml) was transferred to a 96 well plate and the OD at 600 nm was measured (Fluostar Galaxy plate reader BMG), the remaining 750 µl centrifuged down. Next the pellet was re-suspended in S-complete fed (S-complete supplemented with 0.1mg/ml ampiciline and 1 mM IPTG) and volume adjusted to obtain OD₆₀₀ value of 1

Example 2: Generation of fat storage phenotype in *C. elegans*.

In this example, *C. elegans* strain CB1370 containing the temperature sensitive *daf-2* allele e-1370 is used (Ogg et al., supra). CB 1370 is publicly available from, for example, the Caenorhabditis Genetics Center (CGC), Minnesota, USA).

To generate the fat-storage phenotype, L1 worms of strain CB 1370 were cultivated at a temperature of 15 °C in S-Complete fed-medium in the wells of a 96 well plate (1 L1 nematodes per well) under essentially synchronized conditions, until the nematodes reached the L2 stage.

Then, the temperature was increased to 20°C, and the worms were further cultivated at said temperature until their F1 offspring reached the L4 stage (about 144 hrs). Due to the presence of the *daf-2* allele e-1370, this raise in temperature from 15°C to 20°C causes the nematodes to accumulate fat, mainly in their intestinal and hypodermal tissue (vide Ogg et al. and Figures 2A and 2B).

The accumulation of fat (in the form of droplets) was made visible by means of Nile Red staining: L4 animals were washed several times with M9 (supplemented with 0.1% PEG) to remove the remaining *E.coli*, and fixed with MeOH (fc. 33%). After fixation the nematodes were stained with nile red (fc 0.375 mM in 37.5% MeOH) for 4 hours. MeOH and excess dye was removed through several washes with M9 (supplemented with 0.1% PEG). The staining pattern was visualized under UV using a 500 nm long pass filter.

For testing the influence of the gene T17E9.1a on fat storage, during the steps described above, the worms were grown on 30 µl of the normalized *E.coli* pGN49A-T17E9.1a strain of Example 1 (OD₆₀₀ = 1) as a food source. As a reference, the daf-2 (e1370) nematodes were grown in a similar manner, but with *E. coli* strain AB 309-105 containing vector pGN29A without the T17E9.1a insert as a food source, used in the same amount. All samples were carried out in quadruplicate.

The results were as follows: worms fed on *E.coli* pGN49A-T17E9.1a strain, which downregulates the expression of T17E9.1a through RNA interference, showed a strong reduction of the accumulation of fat, compared to the reference (vide Figures 2A and 2B).

These results show that T17E9.1a is involved in the regulation of (the daf-2 dependent) the *daf*-pathway and the accumulation of fat in the nematode. It is known in the art that both are models for insulin resistance and fat handling in mammals, such as humans.

Bio-informatic analysis of the sequence of T17E9.1a (kin-18) showed that the amino acid sequence of the kin-18 protein (SEQ ID NO:2) has the following degree of sequence identity (in percentages at the amino acid level) with the human amino acid sequences of SEQ ID NOS: 4, 6, 8 and 10, respectively:

- SEQ ID NO:4 ("JIK"): overall sequence: 36%; kinase domain: 61%
- SEQ ID NO:6 ("PSK"): overall sequence: 24%; kinase domain: 59%
- SEQ ID NO:8 ("TAOI"): overall sequence: 32%; kinase domain: 59%
- SEQ ID NO:10 ("Q9P2I6"): overall sequence: 33%; kinase domain: 61%

SEQUENCE LISTING

5 <110> Devgen N.V.

Devgen N.V.

10

<120> Kinase sequences useful for developing compounds for the prevention and/or treatment of metabolic diseases and nucleotide sequences encoding such kinase sequences.

15

<130> P 02-011

20

<160> 13

25

<170> PatentIn version 3.1

30

<210> 1

<211> 2949

<212> DNA

35

<213> Caenorhabditis elegans

40

<400> 1

atggcgccctg ccgtcttaca aaaacccggt gttatcaagg atccatcgat tgctgcattg
60

ttcagtaata aggatccaga gcagagatat caagatttaa gagaaattgg acatggatct
120

45

tttggagctg tctattttgc atatgacaaa aaaaatgagc agactgttgc gattaataag
180

50

atgaatttta gtggaaaaca ggctgtcgaa aaatggaatg atattcttaa agaagtgtct
240

tttctgaata cagttgttca tccacatatt gtcgactaca aggcttgttt tcttaaggac
300

55

actacatgtt ggcttgtgat ggagtactgt attggctctg cagccgatat agtggatgtc
360

ttgcgaaaag gaatgcgaga agtcgaaatc gctgcgattt gctctcaaac tttggatgct.
 420
 5 cttcgatata ttactctct gaagcgaata catcgagata ttaaagctgg aaatattctg
 480
 ctatctgac atgctattgt taaactagct gatttcggat ccgcatccct ggtagatccg
 540
 10 gctcaaactt tcatcggaac gccgtttttc atggccccag aggtaattct ggcaatggat
 600
 gagggtcact acacggatcg tgcagatatt tggtcattgg gtatcacgtg tatagagctg
 660
 15 gccgaacgtc gtccaccatt gttcagtatg aatgcaatgt ctgccctcta ccatattgct
 720
 caaatgata ctccaactct ttctccaatt gacactagcg aacaaccgga atggtcgctg
 20 780
 gaattcgctt aatttataga caaatgtctt cgaaaaccag cagaagagcg aatgtcagct
 840
 25 gaagaatgct ttcgacatcc attcattcaa cggctctgcc catcagacac aattcaggaa
 900
 ctattcaga gaacgaaaaa tatggtatta gagttggata attttcaata caaaaagatg
 960
 30 agaaaactca tgtatttggg tgaaacagaa ggaaaagaag gaagtgaagg aaatggagca
 1020
 tctgatgatt tagattttca tggaaatgaa gctaattcaa ttggaagagc aggagattct
 35 1080
 gcgtcatctc gaagtgttct tottacttct ttccgatcaa tgcagagtag tggaggagct
 1140
 40 ggtcttttag tgtccaccaa tacgacgggt gctatggata atgtgcatgg atcctctgga
 1200
 tacggtaatg gaagtagttc gacgacgagc tccgcacgcc gccgtcctcc aattccttcg
 1260
 45 caaatgtctt cttctacatc aacgtctggt gttggaacta tgccgagtca tggatcagtt
 1320
 ggagcatcga ttacggcgat cgcagtcaat ccaacaccgt ctcttcaga acctatccca
 50 1380
 acatcacaac caacatcgaa atcagaatca tttctatac tcgaaactgc acacgatgat
 1440
 55 cctttggaca cgtcgatacg tgctccagtg aaagaattgc atatgccgca tcgagcagtc
 1500

aaggaacgaa tagccacggt gcaaaatcac aaattcgga cgcttcgttc ccagagaata
1560

5 atcaatcagg aacaagaaga atatacgaag gagaacaata tgtatgagca aatgagcaag
1620

tacaagcatc tacgacaagc acatcacaaa gagctccaac aatttgaaga acgatgtgca
1680

10 ttagatagag agcaactgag tgtgaaaatg gatcgagaac tcgaacaatt gacaacgaca
1740

tactcgaaag aaaagatgag agtgagggtg tcacagaata atgaactaga caaacggaaa
1800

15 aaagatatcg aagatgggga gaaaaagatg aaaaagacga aaaatagtca aaatcagcag
1860

20 cagatgaaac tgtattcagc gcaacaattg aaagaatata agtataacaa ggaggcacag
1920

aaaacacgat tacgaagtct gaacatgcct cgaagtactt atgagaacgc aatgaaagaa
1980

25 gtgaaagccg atctgaatcg agtgaaagat gcacgggaaa atgattttga cgagaagctt
2040

cgtgcagaac ttgaagatga aattgtaagg tatcgaggc aacaactcag taatcttcat
2100

30 caattggaag aacaattgga tgatgaagac gtaaactgac aagaacgcca aatggacacg
2160

35 cgtcacgat tactgtcaaa gcagcatgaa atgacgcgcg atttggaat acagcatctc
2220

aacgagcttc acgcgatgaa aaaacgacat ttggagacac aacacgaggc ggaatcggca
2280

40 agtcaaaatg agtacacaca gaggcaacag gatgaattga gaaaaaagca tgcgatgcag
2340

tcaagacaac agccaagaga tttaaagatc caagaagcac aaattcgaaa acaataccga
2400

45 caagttgtga agactcagac tcgccaattt aagctctacc ttacacaaat ggtgcaagta
2460

50 gttcctaaaag atgaacaaaa agagctcagc tctcgactaa aacaggatca aatgcaaaaa
2520

gtcgcacttc ttgcttcaca atacgaaagt caaatcaaaa aaatggttca ggataagaca
2580

55 gtgaagctcg agtcgtggca agaagatgaa caacgggttc ttagtgagaa gttggagaaa
2640

gaattggaag aattgattgc ttatcagaag aagacgagag ccacattaga agagcagatt
2700

5 aaaaaggaac gtacggcact cgaagaacga attggcacac gacgtgcaat gcttgaacag
2760

aagattattg aagaacgcga acaaattggga gaaatgcgtc gactaaagaa ggagcaaattc
2820

10 cgtgatcgac acagtcaaga acgccatcgt ctcgagaatc atttcgtacg gacgggctcg
2880

acgagcagaa gttctgggtgg gatcgctcct ggtgttgga attcaagcag tattcagatg
2940

15 gctatgtag
2949

20 <210> 2

<211> 982

<212> PRT

25 <213> Caenorhabditis elegans

30 <400> 2

Met Ala Pro Ala Val Leu Gln Lys Pro Gly Val Ile Lys Asp Pro Ser
1 5 10 15

35 Ile Ala Ala Leu Phe Ser Asn Lys Asp Pro Glu Gln Arg Tyr Gln Asp
20 25 30

40 Leu Arg Glu Ile Gly His Gly Ser Phe Gly Ala Val Tyr Phe Ala Tyr
35 40 45

45 Asp Lys Lys Asn Glu Gln Thr Val Ala Ile Lys Lys Met Asn Phe Ser
50 55 60

Gly Lys Gln Ala Val Glu Lys Trp Asn Asp Ile Leu Lys Glu Val Ser
65 70 75 80

50 Phe Leu Asn Thr Val Val His Pro His Ile Val Asp Tyr Lys Ala Cys
85 90 95

55 Phe Leu Lys Asp Thr Thr Cys Trp Leu Val Met Glu Tyr Cys Ile Gly
100 105 110

5 Ser Ala Ala Asp Ile Val Asp Val Leu Arg Lys Gly Met Arg Glu Val
 115 120 125
 10 Glu Ile Ala Ala Ile Cys Ser Gln Thr Leu Asp Ala Leu Arg Tyr Leu
 130 135 140
 15 His Ser Leu Lys Arg Ile His Arg Asp Ile Lys Ala Gly Asn Ile Leu
 145 150 155 160
 20 Leu Val Asp Pro Ala Gln Thr Phe Ile Gly Thr Pro Phe Phe Met Ala
 180 185 190
 25 Pro Glu Val Ile Leu Ala Met Asp Glu Gly His Tyr Thr Asp Arg Ala
 195 200 205
 30 Asp Ile Trp Ser Leu Gly Ile Thr Cys Ile Glu Leu Ala Glu Arg Arg
 210 215 220
 35 Pro Pro Leu Phe Ser Met Asn Ala Met Ser Ala Leu Tyr His Ile Ala
 225 230 235 240
 40 Gln Asn Asp Pro Pro Thr Leu Ser Pro Ile Asp Thr Ser Glu Gln Pro
 245 250 255
 45 Glu Trp Ser Leu Glu Phe Val Gln Phe Ile Asp Lys Cys Leu Arg Lys
 260 265 270
 50 Pro Ala Glu Glu Arg Met Ser Ala Glu Glu Cys Phe Arg His Pro Phe
 275 280 285
 55 Ile Gln Arg Ser Arg Pro Ser Asp Thr Ile Gln Glu Leu Ile Gln Arg
 290 295 300
 60 Thr Lys Asn Met Val Leu Glu Leu Asp Asn Phe Gln Tyr Lys Lys Met
 305 310 315 320
 65 Arg Lys Leu Met Tyr Leu Asp Glu Thr Glu Gly Lys Glu Gly Ser Glu
 325 330 335

Gly Asn Gly Ala Ser Asp Asp Leu Asp Phe His Gly Asn Glu Ala Asn
 340 345 350

5 Ser Ile Gly Arg Ala Gly Asp Ser Ala Ser Ser Arg Ser Ala Ser Leu
 355 360 365

10 Thr Ser Phe Arg Ser Met Gln Ser Ser Gly Gly Ala Gly Leu Leu Val
 370 375 380

15 Ser Thr Asn Thr Thr Gly Ala Met Asp Asn Val His Gly Ser Ser Gly
 385 390 395 400

20 Tyr Gly Asn Gly Ser Ser Ser Thr Thr Ser Ser Ala Arg Arg Arg Pro
 405 410 415

Pro Ile Pro Ser Gln Met Leu Ser Ser Thr Ser Thr Ser Gly Val Gly
 420 425 430

25 Thr Met Pro Ser His Gly Ser Val Gly Ala Ser Ile Thr Ala Ile Ala
 435 440 445

30 Val Asn Pro Thr Pro Ser Pro Ser Glu Pro Ile Pro Thr Ser Gln Pro
 450 455 460

35 Thr Ser Lys Ser Glu Ser Ser Ser Ile Leu Glu Thr Ala His Asp Asp
 465 470 475 480

Pro Leu Asp Thr Ser Ile Arg Ala Pro Val Lys Asp Leu His Met Pro
 485 490 495

40 His Arg Ala Val Lys Glu Arg Ile Ala Thr Leu Gln Asn His Lys Phe
 500 505 510

45 Ala Thr Leu Arg Ser Gln Arg Ile Ile Asn Gln Glu Gln Glu Glu Tyr
 515 520 525

50 Thr Lys Glu Asn Asn Met Tyr Glu Gln Met Ser Lys Tyr Lys His Leu
 530 535 540

55 Arg Gln Ala His His Lys Glu Leu Gln Gln Phe Glu Glu Arg Cys Ala
 545 550 555 560

Leu Asp Arg Glu Gln Leu Arg Val Lys Met Asp Arg Glu Leu Glu Gln
 565 570 575

5 Leu Thr Thr Thr Tyr Ser Lys Glu Lys Met Arg Val Arg Cys Ser Gln
 580 585 590

10 Asn Asn Glu Leu Asp Lys Arg Lys Lys Asp Ile Glu Asp Gly Glu Lys
 595 600 605

15 Lys Met Lys Lys Thr Lys Asn Ser Gln Asn Gln Gln Gln Met Lys Leu
 610 615 620

Tyr Ser Ala Gln Gln Leu Lys Glu Tyr Lys Tyr Asn Lys Glu Ala Gln
 625 630 635 640

20 Lys Thr Arg Leu Arg Ser Leu Asn Met Pro Arg Ser Thr Tyr Glu Asn
 645 650 655

25 Ala Met Lys Glu Val Lys Ala Asp Leu Asn Arg Val Lys Asp Ala Arg
 660 665 670

30 Glu Asn Asp Phe Asp Glu Lys Leu Arg Ala Glu Leu Glu Asp Glu Ile
 675 680 685

35 Val Arg Tyr Arg Arg Gln Gln Leu Ser Asn Leu His Gln Leu Glu Glu
 690 695 700

Gln Leu Asp Asp Glu Asp Val Asn Val Gln Glu Arg Gln Met Asp Thr
 705 710 715 720

40 Arg His Gly Leu Leu Ser Lys Gln His Glu Met Thr Arg Asp Leu Glu
 725 730 735

45 Ile Gln His Leu Asn Glu Leu His Ala Met Lys Lys Arg His Leu Glu
 740 745 750

50 Thr Gln His Glu Ala Glu Ser Ala Ser Gln Asn Glu Tyr Thr Gln Arg
 755 760 765

55 Gln Gln Asp Glu Leu Arg Lys Lys His Ala Met Gln Ser Arg Gln Gln
 770 775 780

Pro Arg Asp Leu Lys Ile Gln Glu Ala Gln Ile Arg Lys Gln Tyr Arg

	785	790	795	800
5	Gln Val Val Lys Thr	Gln Thr Arg Gln Phe Lys Leu Tyr Leu Thr Gln		
	805	810		815
10	Met Val Gln Val Val Pro Lys Asp Glu Gln Lys Glu Leu Thr Ser Arg			
	820	825		830
15	Leu Lys Gln Asp Gln Met Gln Lys Val Ala Leu Leu Ala Ser Gln Tyr			
	835	840		845
20	Glu Ser Gln Ile Lys Lys Met Val Gln Asp Lys Thr Val Lys Leu Glu			
	850	855		860
25	Ser Trp Gln Glu Asp Glu Gln Arg Val Leu Ser Glu Lys Leu Glu Lys			
	865	870		875
30	Glu Leu Glu Glu Leu Ile Ala Tyr Gln Lys Lys Thr Arg Ala Thr Leu			
		885	890	895
35	Glu Glu Gln Ile Lys Lys Glu Arg Thr Ala Leu Glu Glu Arg Ile Gly			
		900	905	910
40	Thr Arg Arg Ala Met Leu Glu Gln Lys Ile Ile Glu Glu Arg Glu Gln			
		915	920	925
45	Met Gly Glu Met Arg Arg Leu Lys Lys Glu Gln Ile Arg Asp Arg His			
		930	935	940
50	Ser Gln Glu Arg His Arg Leu Glu Asn His Phe Val Arg Thr Gly Ser			
		945	950	955
55	Thr Ser Arg Ser Ser Gly Gly Ile Ala Pro Gly Val Gly Asn Ser Ser			
		965	970	975
	Ser Ile Gln Met Ala Met			
		980		
	<210> 3			
	<211> 4188			
	<212> DNA			

<213> Homo sapiens

5 <300>
 <308> genbank
 <309> 2002-10-04
 10 <313> (1)..(4188)

15 <300>
 <308> genbank NM_016281
 <309> 2002-10-04
 20 <313> (1)..(4188)

25 <400> 3
 gccggggaac aagccacagg agagcgactc aggaacaagt gtgggagagg aagcggcggc
 60
 ggcggcgcgc ggcccggggg tggcgacagc aggtctgagg ttgcatcata aatacaaagg
 30 120
 actgaagtta taaaagagaa aagagaagtt tgctgctaaa atgaatctga gcaatatgga
 180

35 atattttgtg ccacacacaa aaaggtactg aagatttacc ccccaaaaaa aattgtcaat
 240
 gagaaataaa gctaactgat atcaaaaagc agagcctgct ctactggcca tcatgcgtaa
 300
 40 aggggtgctg aaggaccag agattgccga tctattctac aaagatgatc ctgaggaact
 360

45 ttttattggt ttgcatgaaa ttggacatgg aagttttgga gcagtttatt ttgctacaaa
 420
 tgctcacacc aatgaggtgg tggcaattaa gaagatgtcc tatagtggga agcagaccca
 480

50 tgagaaatgg caagatatcc ttaaggaagt taaattttta cgacaattga agcatcctaa
 540
 tactattgag tacaaaggct gttacttgaa agaacacact gottgggttg tgatggaata
 600
 55 ttgcttaggc tcagcctctg atttattaga agttcataaa aaaccacttc aggaagtgga
 660

gatcgctgcc attactcatg gagccttgca tggactagcc tacctacatt ctcatgcatt
720

5 gattcatagg gatattaaag caggaaatat tcttctaaca gagccaggtc aggtaaaact
780

agctgatttt ggatctgctt caatggcttc tcttgccaac tccttcgtgg gcacacctta
840

10 ctggatggct ccagagggtga tcttagctat ggatgaagga cagtatgatg ggaaagttga
900

15 tatttgggtca cttggcatca cttgtattga attggcggaa cggaagccgc cccttttcaa
960

catgaatgca atgagtgcct tatatcacat tgcccagaat gactcccca cgttacagtc
1020

20 taatgaatgg acagactcct ttaggagatt tggtgattac tgcttgacaga aaataacctca
1080

ggaaaggcca acatcagcag aactattaag gcatgacttt gttcgacgag accggccact
1140

25 acgtgtcctc attgacctca tacagaggac aaaagatgca gttcgtgagc tagataacct
1200

30 acagtaccga aaaatgaaaa aaatactttt ccaagagaca cggaatggac ccttgaatga
1260

gtcacaggag gatgagggaag acagtgaaca tggaaccagc ctgaacaggg aaatggacag
1320

35 cctgggcagc aaccattcca ttccaagcat gtccgtgagc acaggcagcc agagcagcag
1380

tgtgaacagc atgcaggaag tcatggacga gagcagttcc gaacttgtca tgatgcacga
1440

40 tgacgaaagc acaatcaatt ccagctcctc cgtcgtgcat aagaagatc atgtattcat
1500

45 aagggatgag gcggggccacg gcatcccag gcctgagccg cggcctaccc agtcagttca
1560

gagccaggcc ctccactacc ggaacagaga gcgctttgcc acgatcaa atcagcatcttt
1620

50 ggttacacga cagatccatg agcatgagca ggagaacgag ttgcgggaac agatgtcagg
1680

ttataagcgg atgcggcgcc agcaccagaa gcagctgac gccctggaga acaagctgaa
1740

55 ggctgagatg gacgagcacc gcctcaagct acagaaggag gtggagacgc atgccaacaa
1800

ctcgtccatc gagctggaga agctggccaa gaagcaagtg gctatcatag aaaaggaggc
1860

5 aaaggtagct gcagcagatg agaagaagtt ccagcaacag atcttggccc agcagaagaa
1920

agatttgaca actttcttag aaagtcagaa gaagcagtat aagatttgta aggaaaaaat
1980

10 aaaagaggaa atgaatgagg accatagcac acccaagaaa gagaagcaag agcggatctc
2040

15 caaacataaa gagaacttgc agcacacaca ggctgaagag gaagcccacc ttctcactca
2100

acagagactg tactacgaca aaaattgtcg tttcttcaag cggaaaataa tgatcaagcg
2160

20 gcacgaggtg gagcagcaga acattcggga ggaactaaat aaaaagagga cccagaagga
2220

gatggagcat gccatgctaa tccggcacga cgagtccacc cgagagctag agtacaggca
2280

25 gctgcacacg ttacagaagc tacgcatgga tctgatccgt ttacagcacc agacggaaact
2340

30 ggaaaaccag ctggagtaca ataagaggcg agaaagagaa ctgcacagaa agcatgtcat
2400

ggaacttcgg caacagccaa aaaacttaaa ggccatggaa atgcaaatta aaaaacagtt
2460

35 tcaggacact tgcaaagtac agaccaaaca gtataaagca ctcaagaatc accagttgga
2520

agttactcca aagaatgagc acaaaacaat cttaaagaca ctgaaagatg agcagacaag
2580

40 aaaacttgcc attttggcag agcagtatga acagagtata aatgaaatga tggcctctca
2640

agcgttacgg ctagatgagg ctcaagaagc agaatgccag gcottgaggc tacagctcca
2700

45 gcaggaaatg gagctgctca acgcctacca gagcaaaatc aagatgcaaa cagaggcaca
2760

50 acatgaacgt gagctccaga agctagagca gagagtgtct ctgcgcagag cacaccttga
2820

gcagaagatt gaagaggagc tggctgccct tcagaaggaa cgcagcgaga gaataaagaa
2880

55 cctattggaa aggcaagagc gagagattga aacttttgac atggagagcc tcagaatggg
2940

atttgggaat ttggttacat tagattttcc taaggaggac tacagatgag attaaatttt
3000

5 ttgccattta caaaaaaaaa aaaaaaaga aaacagaaaa aaattcagac cctgcaaaac
3060

cacattcccc attttaacgg gcgttgctct cactctctct ctctcttact cttactgaca
3120

10 tcgtgtcgga ctagtgcctg tttattctta ctccatcagg ggcccccttc ctccccccgt
3180

15 gtcaactttc agtgctggcc aaaacctggc cgtctcttct attcacagta cacgtcacag
3240

tattgatgtg attcaaaatg tttcagtga aactttggag acagttttaa caaaaccaat
3300

20 aaaccaacaa caaaaaaagt ggatgtatat tgctttaagc aatcactcat taccaccaat
3360

ctgtgaaagt aaagcaaaaa ataataataa taaatgccaa gggggagaga gacacaatat
3420

25 ccgcagcctt acaccttaac tagctgctgc attattttat tttattttat ttttttgga
3480

30 tttattcatc aggaataaaaa aaacaaagt tttattaaag attgaaaatt tgatacattt
3540

tacagaaact aattgtgatg tacatatcag tggtgacata ttattacttt tttggggacg
3600

35 ggggtgggtg gggatgaagag atcttgtgat tttagactgc tgcagagtta acttgtctca
3660

gcatatctga tgtatcataa tcatttctgc tgtgcagagg agggatacac ttaggggctc
3720

40 acagatccca gtagcacaat tgggctttgg caaatgggta ttttgtgtat agaggaattt
3780

45 aaggagaggt attacttatt ttcataattgt attttaactg tttctcggat caaatttttt
3840

aacttcttct tcgtgttctt cccacactcc ttccttttcc agttcagtat ttggagtcca
3900

50 acactgtctc tcaatcagat catctggatc tttttcttta tctcccttcc ccttectaag
3960

tcccatttct tggtcataaa tattgcatta ttcacacttt caaactgtgt attttcttac
4020

55 aataaaaaat gatgaaaaaa aaaaaggctt tacttctttt gcatgcactt taaaaacaaa
4080

5 ttgtaaataa aagtgttcat cacaaaaaaaaa aaaaaaaaaa aaaaaaaaaa
4188

<210> 4

10 $\langle 211 \rangle$ 898

<212> PRT

15 <213> Homo sapiens

<300>

20
<308> Genbank NP_057365.2

<309> 2002-10-01

25 <313> (1) . . (898) .

<400> 4

30 Met Arg Lys Gly Val Leu Lys Asp Pro Glu Ile Ala Asp Leu Phe Tyr
1 5 10 15

35 Lys Asp Asp Pro Glu Glu Leu Phe Ile Gly Leu His Glu Ile Gly His
 20 25 30

40 Gly Ser Phe Gly Ala Val Tyr Phe Ala Thr Asn Ala His Thr Asn Glu
35 40 45

45. Val Val Ala Ile Lys Lys Met Ser Tyr Ser Gly Lys Gln Thr His Glu
50 55 60

Lys Trp Gln Asp Ile Leu Lys Glu Val Lys Phe Leu Arg Gln Leu Lys
65 70 75 80

50 His Pro Asn Thr Ile Glu Tyr Lys Gly Cys Tyr Leu Lys Glu His Thr
85 90 95

55 Ala Trp Leu Val Met Glu Tyr Cys Leu Gly Ser Ala Ser Asp Leu Leu
100 105 110

Glu Val His Lys Lys Pro Leu Gln Glu Val Glu Ile Ala Ala Ile Thr
 115 120 125

5 His Gly Ala Leu His Gly Leu Ala Tyr Leu His Ser His Ala Leu Ile
 130 135 140

10 His Arg Asp Ile Lys Ala Gly Asn Ile Leu Leu Thr Glu Pro Gly Gln
 145 150 155 160

15 Val Lys Leu Ala Asp Phe Gly Ser Ala Ser Met Ala Ser Pro Ala Asn
 165 170 175

20 Ser Phe Val Gly Thr Pro Tyr Trp Met Ala Pro Glu Val Ile Leu Ala
 180 185 190

Met Asp Glu Gly Gln Tyr Asp Gly Lys Val Asp Ile Trp Ser Leu Gly
 195 200 205

25 Ile Thr Cys Ile Glu Leu Ala Glu Arg Lys Pro Pro Leu Phe Asn Met
 210 215 220

30 Asn Ala Met Ser Ala Leu Tyr His Ile Ala Gln Asn Asp Ser Pro Thr
 225 230 235 240

35 Leu Gln Ser Asn Glu Trp Thr Asp Ser Phe Arg Arg Phe Val Asp Tyr
 245 250 255

40 Cys Leu Gln Lys Ile Pro Gln Glu Arg Pro Thr Ser Ala Glu Leu Leu
 260 265 270

Arg His Asp Phe Val Arg Arg Asp Arg Pro Leu Arg Val Leu Ile Asp
 275 280 285

45 Leu Ile Gln Arg Thr Lys Asp Ala Val Arg Glu Leu Asp Asn Leu Gln
 290 295 300

50 Tyr Arg Lys Met Lys Lys Ile Leu Phe Gln Glu Thr Arg Asn Gly Pro
 305 310 315 320

55 Leu Asn Glu Ser Gln Glu Asp Glu Glu Asp Ser Glu His Gly Thr Ser
 325 330 335

Leu Asn Arg Glu Met Asp Ser Leu Gly Ser Asn His Ser Ile Pro Ser
 340 345 350

5 Met Ser Val Ser Thr Gly Ser Gln Ser Ser Ser Val Asn Ser Met Gln
 355 360 365

10 Glu Val Met Asp Glu Ser Ser Ser Glu Leu Val Met Met His Asp Asp
 370 375 380

15 Glu Ser Thr Ile Asn Ser Ser Ser Ser Val Val His Lys Lys Asp His
 385 390 395 400

20 Val Phe Ile Arg Asp Glu Ala Gly His Gly Asp Pro Arg Pro Glu Pro
 405 410 415

25 Arg Pro Thr Gln Ser Val Gln Ser Gln Ala Leu His Tyr Arg Asn Arg
 420 425 430

30 Glu Arg Phe Ala Thr Ile Lys Ser Ala Ser Leu Val Thr Arg Gln Ile
 435 440 445

35 His Glu His Glu Gln Glu Asn Glu Leu Arg Glu Gln Met Ser Gly Tyr
 450 455 460

40 Lys Arg Met Arg Arg Gln His Gln Lys Gln Leu Ile Ala Leu Glu Asn
 465 470 475 480

45 Lys Leu Lys Ala Glu Met Asp Glu His Arg Leu Lys Leu Gln Lys Glu
 485 490 495

50 Val Glu Thr His Ala Asn Asn Ser Ser Ile Glu Leu Glu Lys Leu Ala
 500 505 510

55 Lys Lys Gln Val Ala Ile Ile Glu Lys Glu Ala Lys Val Ala Ala Ala
 515 520 525

60 Asp Glu Lys Lys Phe Gln Gln Gln Ile Leu Ala Gln Gln Lys Lys Asp
 530 535 540

65 Leu Thr Thr Phe Leu Glu Ser Gln Lys Lys Gln Tyr Lys Ile Cys Lys
 545 550 555 560

70 Glu Lys Ile Lys Glu Glu Met Asn Glu Asp His Ser Thr Pro Lys Lys

565

570

575

5 Glu Lys Gln Glu Arg Ile Ser Lys His Lys Glu Asn Leu Gln His Thr
 580 585 590

10 Gln Ala Glu Glu Glu Ala His Leu Leu Thr Gln Gln Arg Leu Tyr Tyr
 595 600 605

15 Asp Lys Asn Cys Arg Phe Phe Lys Arg Lys Ile Met Ile Lys Arg His
 610 615 620

20 Glu Val Glu Gln Gln Asn Ile Arg Glu Glu Leu Asn Lys Lys Arg Thr
 625 630 635 640

25 Gln Lys Glu Met Glu His Ala Met Leu Ile Arg His Asp Glu Ser Thr
 645 650 655

30 Arg Glu Leu Glu Tyr Arg Gln Leu His Thr Leu Gln Lys Leu Arg Met
 660 665 670

35 Asp Leu Ile Arg Leu Gln His Gln Thr Glu Leu Glu Asn Gln Leu Glu
 675 680 685

40 Tyr Asn Lys Arg Arg Glu Arg Glu Leu His Arg Lys His Val Met Glu
 690 695 700

45 Leu Arg Gln Gln Pro Lys Asn Leu Lys Ala Met Glu Met Gln Ile Lys
 705 710 715 720

50 Lys Gln Phe Gln Asp Thr Cys Lys Val Gln Thr Lys Gln Tyr Lys Ala
 725 730 735

55 Leu Lys Asn His Gln Leu Glu Val Thr Pro Lys Asn Glu His Lys Thr
 740 745 750

60 Ile Leu Lys Thr Leu Lys Asp Glu Gln Thr Arg Lys Leu Ala Ile Leu
 755 760 765

65 Ala Glu Gln Tyr Glu Gln Ser Ile Asn Glu Met Met Ala Ser Gln Ala
 770 775 780

70 Leu Arg Leu Asp Glu Ala Gln Glu Ala Glu Cys Gln Ala Leu Arg Leu
 785 790 795 800

5 Gln Leu Gln Gln Glu Met Glu Leu Leu Asn Ala Tyr Gln Ser Lys Ile
805 810 815

Lys Met Gln Thr Glu Ala Gln His Glu Arg Glu Leu Gln Lys Leu Glu
820 825 830

10 Gln Arg Val Ser Leu Arg Arg Ala His Leu Glu Gln Lys Ile Glu Glu
835 840 845

15 Glu Leu Ala Ala Leu Gln Lys Glu Arg Ser Glu Arg Ile Lys Asn Leu
850 855 860

20 Leu Glu Arg Gln Glu Arg Glu Ile Glu Thr Phe Asp Met Glu Ser Leu
865 870 875 880

25 Arg Met Gly Phe Gly Asn Leu Val Thr Leu Asp Phe Pro Lys Glu Asp
885 890 895

Tyr Arg

30

<210> 5

<211> 4971

35 <212> DNA

<213> Homo sapiens

40

<300>

<308> Genbank NM_016151.1

45 <309> 2002-11-05

<313> (1)..(4971)

50

<400> 5
aattcggcac gagctgagac ggagaagagg agaggcagag agggcgcgagg gaccgtcagg
60

55 agcaccttag ctacaatcgt tcagctattc tcggaagaga gaaggagag ggaggaggcc
120

ggggcgggag tgggggctgt caccctcgga ccccggcgtg agaggggccc tgccggccgga
180

5 cgctctcggg gtgggcccc agtcggtggc cgaagaccta cagctcaggc cctgggtcc
240

caaatttcca ggctttgccc ctctctctt ctcagatacc cgggtaacag tcctcatagt
300

10 ccagatatcc gggactcggg tcccaacctc tctaaacctg ggtctctgtt tcatagattt
360

tcaaatatca ggttcaggcc cctgcgtgca ccagtatccg gggttcattc cccgggcgtt
420

15 tcaaatatcg gattcagtct ccattccggt cagatattcg gggttcagac cccacaatca
480

20 gaaatccgga attcggcagc tgctgccctc gacgaggggg aggactggac cgcgaggta
540

gattaggttg tcacccccctc cctccaggg gaggtctccc gggcccgccc ctcaggaagg
600

25 gcgaaagccg aggaagaggt ggcaagggga aaggtctcct tgccctctc cctgcttgcc
660

agagccgctg gaggaccca ggcggaagcg gaggcgctgg ggcaccatag tgaccctac
720

30 caggccaggc cccactctca gggccccag gggccaccat gccagctggg ggccgggccc
780

35 ggagcctgaa ggaccagat gtggctgagc tcttcttcaa ggatgaccca gaaaagctct
840

tctctgacct ccgggaaatt ggccatggca gctttggagc cgtatacttt gcccgggatg
900

40 tccggaatag tgagggtgtg gccatcaaga agatgtccta cagtgggaag cagtccaatg
960

agaaatggca agacatcatc aaggagggtc ggttcttaca gaagctccg catcccaaca
1020

45 ccattcagta ccggggctgt tacctgaggg agcacacggc ttggctggta atggagtatt
1080

50 gcctgggctc agcttctgac cttctagaag tgcacaagaa accccttcag gaggtagaga
1140

tgcagctgt gaccacggg gcgcttcagg gcctggcata tctgcactcc cacaacatga
1200

55 tccataggga tgtgaaggct ggaacatcc tgctgtcaga gccagggtta gtgaagctag
1260

gggacttttg ttctgctcc atcatggcac ctgccaaactc cttcgtgggc accccataact
1320

5 ggatggcacc cgaggtgatc ctggccatgg atgaggggca gtacgatggc aaagtggacg
1380

tctggtcctt ggggataacc tgcacgagc tggctgaacg gaaaccaccg ctctttaaca
1440

10 tgaatgcat gatgcctta taccacattg cacagaacga atccccctg ctccagtcag
1500

gacactggc tgagtacttc cggaattttg tcgactcctg tcttcagaaa atccctcaag
1560

15 acagaccaac ctccagaggtt ctccctgaagc accgctttgt gctccgggag cggccacca
1620

20 cagtcacat ggacctgatc cagaggacca aggatgccgt gcgggagctg gacaacctgc
1680

agtaccgcaa gatgaagaag atcctgttcc aagaggcacc caacggccct ggtgccgagg
1740

25 cccagagga ggaagaggag gccgagccct acatgcaccg ggccgggact ctgaccagcc
1800

tcgagagtag ccactcagt cccagcatgt ccatcagcgc ctccagccag agcagctccg
1860

30 tcaacagcct agcagatgcc tcagacaacg aggaagagga ggaggaggag gaggaagagg
1920

35 aggaggagga agaaggccct gaagcccggg agatggccat gatgcaggag ggggagcaca
1980

cagtcacctc tcacagctcc attatccacc ggctgccggg ctctgacaac ctatatgatg
2040

40 acccctacca gccagagata acccccagcc ctctccagcc gctgcagcc ccagctccca
2100

cttcaccac ctcttcgcc cgccgccggg cctactgccg taaccgagac cactttgcca
2160

45 ccatccgaac cgcctccctg gtcagccgtc agatccagga gcatgagcag gactctgcgc
2220

tgccggagca gctgagcggc tataagcggg tgccagcaca gcaccagaag cagctgctgg
2280

50 ccctggagtc acggctgagg ggtgaacggg aggagcacag tgcacggctg cagcgggagc
2340

55 ttgaggcgca gcgggctggc tttggggcag aggcagaaaa gctggcccgg cggcaccagg
2400

ccataggtga gaaggaggca cgagctgccc aggccgagga gcggaagttc cagcagcaca
2460

5 tccttgggca gcagaagaag gagctggctg ccctgctgga ggcacagaag cggacctaca
2520

aacttcgcaa ggaacagctg aaggaggagc tccaggagaa cccagcact cccaagcggg
2580

10 agaaggccga gtggctgctg cggcagaagg agcagctcca gcagtgccag gcggaggagg
2640

aagcagggt gctgcggcgg cagcgccagt actttgagct gcagtgtcgc cagtacaagc
2700

15 gcaagatgtt gctggctcgg cacagcctgg accaggacct gctgcgggag gacctgaaca
2760

20 agaagcagac ccagaaggac ttggagtgtg cactgctgct tcggcagcac gaggccacgc
2820

gggagctgga gctgcggcag ctccaggccg tgcagcgcac gcgggctgag ctcacccgcc
2880

25 tgcagcacca gacggagctg ggcaaccagc tggagtacaa caagcggcgt gagcaagagt
2940

tgccgcagaa gcatgcggcc caggttcgcc agcagcccaa ggcctcaaa gtacgtgcag
3000

30 gccagcgccc cccgggcctt ccactcccca ttcctggggc tctgggcca cccaacacag
3060

gcacccctat agaacagcag cctgctcac ctggccagga ggcagtcctg gaccaaagaa
3120

35 tgcttggcga ggaggaggaa gcagttggag agagaaggat tctgggaaag gaaggggcca
3180

40 ctttggagcc caagcagcag aggattctgg gggaagaatc aggagcccct agtcccagtc
3240

cacaaaaaca tgggagcctg gttgatgagg aagtttgggg tctgcctgag gagatagagg
3300

45 agcttagggt gccctccctt gtaccccagg agaggagcat tgttgccag gaggaggctg
3360

50 ggacgtggag cttgtggggg aaggaggatg agagtcttct ggatgaggag tttgagcttg
3420

gctgggtcca gggcccagca ctgactcccg tccctgagga ggaggaagaa gaggaagagg
3480

55 gggctccgat tgggaccctc agggatcctg gagatggttg tccttcccc gacatccctc
3540

ctgaaccccc tccaacacac ctgaggccct gccctgccag ccagctccct ggactcctgt
3600

5 cccatggcct cctggccggc ctctcctttg cagtggggtc ctctctggc ctctgcccc
3660

tcttgctgct gctgctgctt ccattgctgg cagcccaggg tgggggtggc ctgcaggcag
3720

10 cgctgctggc ccttgagggtg gggctgggtg gtctgggggc ctctacctg ctccctttgta
3780

cagccctgca cctgccctcc agtcttttcc tactcctggc ccagggtacc gcaactggggg
3840

15 ccgtcctggg cctgagctgg cgccgaggcc tcatgggtgt tcccctgggc cttggagctg
3900

20 cctggctctt agcttggcca ggcctagctc tacctctggt ggctatggca gcggggggca
3960

gatgggtgcg gcagcagggc ccccggtg cccggggcat atctcgactc tggttgcggg
4020

25 ttctgctgcg cctgtcacc atggccttcc gggccctgca gggctgtggg gctgtggggg
4080

accgggggtct gtttgcaactg taccacaaaa ccaacaagga tggcttccgc agccgcctgc
4140

30 ccgtccctgg gccccggcgg cgtaatcccc gcaccacca acaccatta gctctgttgg
4200

35 caagggtctg ggtcctgtgc aagggtgga actggcgtct ggcacgggac agccagggtt
4260

tagcatcca cttgccccg tgggccatcc acacactggc cagctggggc ctgcttcggg
4320

40 gtgaacggcc caccgaatc ccccggtac taccacgag ccagcgccag ctagggcccc
4380

ctgcctcca ccagccactg ccagggactc tagccggg cgggtcacgc acccgccagt
4440

45 cccgggccct gccccctgg aggtagctga ctccagccct tccagcccaa atctagagca
4500

50 ttgagcactt tatctccac gactcagtga agtttctcca gtccctagtc ctctcttttc
4560

accaccttc ctcagtttgc tcaattacc caggcccagc cttcgggacc tctagacagg
4620

55 cagcctctc agctgtggag tccagcagtc actctgtgtt ctctgggc tctcccta
4680

agttattgct gttcgcccg c tgtgtgtgct catcctcacc ctcattgact caggcctggg
4740

gccaggggtg gtggaggggtg ggaagagtca tgtttttttt ctcctctttg attttgtttt
5 4800

tctgtctccc ttccaacctg tccccttccc cccaccaaaa aaagaaaaag acaaacacaa
4860

10 ataaaatatc tgagcggaac tgtgaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa
4920

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa a
4971

15

<210> 6

<211> 1235

20

<212> PRT

<213> Homo sapiens

25

<300>

<308> Genbank NP_057235.1

30

<309> 2002-11-05

<313> (1)..(1235)

35

<400> 6

Met Pro Ala Gly Gly Arg Ala Gly Ser Leu Lys Asp Pro Asp Val Ala
40 1 5 10 15

Glu Leu Phe Phe Lys Asp Asp Pro Glu Lys Leu Phe Ser Asp Leu Arg
45 20 25 30

Glu Ile Gly His Gly Ser Phe Gly Ala Val Tyr Phe Ala Arg Asp Val
35 40 45

50

Arg Asn Ser Glu Val Val Ala Ile Lys Lys Met Ser Tyr Ser Gly Lys
50 55 60

55

Gln Ser Asn Glu Lys Trp Gln Asp Ile Ile Lys Glu Val Arg Phe Leu
65 70 75 80

~~Gln-Lys-Leu-Arg-His-Pro-Asn-Thr-Ile-Gln-Tyr-Arg-Gly-Cys-Tyr-Leu~~
85 90 95

5 Arg Glu His Thr Ala Trp Leu Val Met Glu Tyr Cys Leu Gly Ser Ala
100 105 110

10 Ser Asp Leu Leu Glu Val His Lys Lys Pro Leu Gln Glu Val Glu Ile
115 120 125

15 Ala Ala Val Thr His Gly Ala Leu Gln Gly Leu Ala Tyr Leu His Ser
130 135 140

20 His Asn Met Ile His Arg Asp Val Lys Ala Gly Asn Ile Leu Leu Ser
145 150 155 160

25 Glu Pro Gly Leu Val Lys Leu Gly Asp Phe Gly Ser Ala Ser Ile Met
165 170 175

30 Ala Pro Ala Asn Ser Phe Val Gly Thr Pro Tyr Trp Met Ala Pro Glu
180 185 190

35 Val Ile Leu Ala Met Asp Glu Gly Gln Tyr Asp Gly Lys Val Asp Val
195 200 205

40 Trp Ser Leu Gly Ile Thr Cys Ile Glu Leu Ala Glu Arg Lys Pro Pro
210 215 220

45 Leu Phe Asn Met Asn Ala Met Ser Ala Leu Tyr His Ile Ala Gln Asn
225 230 235 240

50 Glu Ser Pro Val Leu Gln Ser Gly His Trp Ser Glu Tyr Phe Arg Asn
245 250 255

55 Phe Val Asp Ser Cys Leu Gln Lys Ile Pro Gln Asp Arg Pro Thr Ser
260 265 270

60 Glu Val Leu Leu Lys His Arg Phe Val Leu Arg Glu Arg Pro Pro Thr
275 280 285

65 Val Ile Met Asp Leu Ile Gln Arg Thr Lys Asp Ala Val Arg Glu Leu
290 295 300

Asp Asn Leu Gln Tyr Arg Lys Met Lys Lys Ile Leu Phe Gln Glu Ala
 305 310 315 320

5 Pro Asn Gly Pro Gly Ala Glu Ala Pro Glu Glu Glu Glu Glu Ala Glu
 325 330 335

10 Pro Tyr Met His Arg Ala Gly Thr Leu Thr Ser Leu Glu Ser Ser His
 340 345 350

15 Ser Val Pro Ser Met Ser Ile Ser Ala Ser Ser Gln Ser Ser Ser Val
 355 360 365

Asn Ser Leu Ala Asp Ala Ser Asp Asn Glu Glu Glu Glu Glu Glu
 370 375 380

20 Glu Glu Glu Glu Glu Glu Glu Glu Gly Pro Glu Ala Arg Glu Met Ala
 385 390 395 400

25 Met Met Gln Glu Gly Glu His Thr Val Thr Ser His Ser Ser Ile Ile
 405 410 415

30 His Arg Leu Pro Gly Ser Asp Asn Leu Tyr Asp Asp Pro Tyr Gln Pro
 420 425 430

35 Glu Ile Thr Pro Ser Pro Leu Gln Pro Pro Ala Ala Pro Ala Pro Thr
 435 440 445

Ser Thr Thr Ser Ser Ala Arg Arg Arg Ala Tyr Cys Arg Asn Arg Asp
 450 455 460

40 His Phe Ala Thr Ile Arg Thr Ala Ser Leu Val Ser Arg Gln Ile Gln
 465 470 475 480

45 Glu His Glu Gln Asp Ser Ala Leu Arg Glu Gln Leu Ser Gly Tyr Lys
 485 490 495

50 Arg Met Arg Arg Gln His Gln Lys Gln Leu Leu Ala Leu Glu Ser Arg
 500 505 510

Leu Arg Gly Glu Arg Glu Glu His Ser Ala Arg Leu Gln Arg Glu Leu
 515 520 525

55

Glu Ala Gln Arg Ala Gly Phe Gly Ala Glu Ala Glu Lys Leu Ala Arg

530

535

540

5 Arg His Gln Ala Ile Gly Glu Lys Glu Ala Arg Ala Ala Gln Ala Glu
 545 550 555 560

10 Glu Arg Lys Phe Gln Gln His Ile Leu Gly Gln Gln Lys Lys Glu Leu
 565 570 575

15 Ala Ala Leu Leu Glu Ala Gln Lys Arg Thr Tyr Lys Leu Arg Lys Glu
 580 585 590

20 Gln Leu Lys Glu Glu Leu Gln Glu Asn Pro Ser Thr Pro Lys Arg Glu
 595 600 605

25 Lys Ala Glu Trp Leu Leu Arg Gln Lys Glu Gln Leu Gln Gln Cys Gln
 610 615 620

30 Ala Glu Glu Glu Ala Gly Leu Leu Arg Arg Gln Arg Gln Tyr Phe Glu
 625 630 635 640

35 Leu Gln Cys Arg Gln Tyr Lys Arg Lys Met Leu Leu Ala Arg His Ser
 645 650 655

40 Leu Asp Gln Asp Leu Leu Arg Glu Asp Leu Asn Lys Lys Gln Thr Gln
 660 665 670

45 Lys Asp Leu Glu Cys Ala Leu Leu Leu Arg Gln His Glu Ala Thr Arg
 675 680 685

50 Glu Leu Glu Leu Arg Gln Leu Gln Ala Val Gln Arg Thr Arg Ala Glu
 690 695 700

55 Leu Thr Arg Leu Gln His Gln Thr Glu Leu Gly Asn Gln Leu Glu Tyr
 705 710 715 720

60 Asn Lys Arg Arg Glu Gln Glu Leu Arg Gln Lys His Ala Ala Gln Val
 725 730 735

65 Arg Gln Gln Pro Lys Ser Leu Lys Val Arg Ala Gly Gln Arg Pro Pro
 740 745 750

70 Gly Leu Pro Leu Pro Ile Pro Gly Ala Leu Gly Pro Pro Asn Thr Gly
 755 760 765

5 Thr Pro Ile Glu Gln Gln Pro Cys Ser Pro Gly Gln Glu Ala Val Leu
 770 775 780

10 Asp Gln Arg Met Leu Gly Glu Glu Glu Glu Ala Val Gly Glu Arg Arg
 785 790 795 800

15 Ile Leu Gly Lys Glu Gly Ala Thr Leu Glu Pro Lys Gln Gln Arg Ile
 805 810 815

20 Leu Gly Glu Glu Ser Gly Ala Pro Ser Pro Ser Pro Gln Lys His Gly
 820 825 830

25 Ser Leu Val Asp Glu Glu Val Trp Gly Leu Pro Glu Glu Ile Glu Glu
 835 840 845

30 Leu Arg Val Pro Ser Leu Val Pro Gln Glu Arg Ser Ile Val Gly Gln
 850 855 860

35 Glu Glu Ala Gly Thr Trp Ser Leu Trp Gly Lys Glu Asp Glu Ser Leu
 865 870 875 880

40 Leu Asp Glu Glu Phe Glu Leu Gly Trp Val Gln Gly Pro Ala Leu Thr
 885 890 895

45 Pro Val Pro Glu Glu Glu Glu Glu Glu Glu Glu Gly Ala Pro Ile Gly
 900 905 910

50 Thr Pro Arg Asp Pro Gly Asp Gly Cys Pro Ser Pro Asp Ile Pro Pro
 915 920 925

55 Glu Pro Pro Pro Thr His Leu Arg Pro Cys Pro Ala Ser Gln Leu Pro
 930 935 940

60 Gly Leu Leu Ser His Gly Leu Leu Ala Gly Leu Ser Phe Ala Val Gly
 945 950 955 960

65 Ser Ser Ser Gly Leu Leu Pro Leu Leu Leu Leu Leu Leu Leu Pro Leu
 965 970 975

70 Leu Ala Ala Gln Gly Gly Gly Gly Leu Gln Ala Ala Leu Leu Ala Leu
 980 985 990

~~Glu Val Gly Leu Val Gly Leu Gly Ala Ser Tyr Leu Leu Leu Cys Thr~~
 995 1000 1005

5

Ala Leu His Leu Pro Ser Ser Leu Phe Leu Leu Leu Ala Gln Gly
 1010 1015 1020

10

Thr Ala Leu Gly Ala Val Leu Gly Leu Ser Trp Arg Arg Gly Leu
 1025 1030 1035

15

Met Gly Val Pro Leu Gly Leu Gly Ala Ala Trp Leu Leu Ala Trp
 1040 1045 1050

20

Pro Gly Leu Ala Leu Pro Leu Val Ala Met Ala Ala Gly Gly Arg
 1055 1060 1065

Trp Val Arg Gln Gln Gly Pro Arg Val Arg Arg Gly Ile Ser Arg
 1070 1075 1080

25

Leu Trp Leu Arg Val Leu Leu Arg Leu Ser Pro Met Ala Phe Arg
 1085 1090 1095

30

Ala Leu Gln Gly Cys Gly Ala Val Gly Asp Arg Gly Leu Phe Ala
 1100 1105 1110

35

Leu Tyr Pro Lys Thr Asn Lys Asp Gly Phe Arg Ser Arg Leu Pro
 1115 1120 1125

40

Val Pro Gly Pro Arg Arg Arg Asn Pro Arg Thr Thr Gln His Pro
 1130 1135 1140

Leu Ala Leu Leu Ala Arg Val Trp Val Leu Cys Lys Gly Trp Asn
 1145 1150 1155

45

Trp Arg Leu Ala Arg Ala Ser Gln Gly Leu Ala Ser His Leu Pro
 1160 1165 1170

50

Pro Trp Ala Ile His Thr Leu Ala Ser Trp Gly Leu Leu Arg Gly
 1175 1180 1185

55

Glu Arg Pro Thr Arg Ile Pro Arg Leu Leu Pro Arg Ser Gln Arg
 1190 1195 1200

Gln Leu Gly Pro Pro Ala Ser His Gln Pro Leu Pro Gly Thr Leu
 1205 1210 1215

5 Ala Gly Arg Arg Ser Arg Thr Arg Gln Ser Arg Ala Leu Pro Pro
 1220 1225 1230

10 Trp Arg
 1235

<210> 7

15 <211> 4242

<212> DNA

<213> Homo sapiens

20

<300>

25 <308> Genbank NM_004783

<309> 2000-11-01

<313> (1)..(4242)

30

<400> 7

agaatttcaa atatcagggt caggccccctg cgtgcaccag tatccgggggt tcattccccg
 60

35

ggcgttcaaa tatcggattc agtctccatc ccgttcagat attcgggggt cagacccac
 120

40

aatcagaaat ccggaattcg gcagctgtcg ccctcgacga gggggaggac tggaccgga
 180

ggtcagatta ggtgtcacc ccctccctc caggggaggc ttcccgggcc cggccctcag
 240

45

gaagggcgaa agccgaggaa gaggtggcaa ggggaaaggt ctcttgccc ctctccctgc
 300

50

ttggcagagc cgctggagga cccagggcgg aagcggaggc gctggggcac catagtgacc
 360

cctaccaggc caggccccac tctcagggcc cccaggggcc accatgccag ctggggggccg
 420

55

ggccggggagc ctgaaggacc cagatgtggc tgagctcttc ttcaaggatg acccagaaaa
 480

gctcttctct gacctccggg aaattggcca tggcagcttt ggagccgtat actttgcccc
540

5 ggatgtccgg aatagtgagg tggtagccat caagaagatg tcctacagtg ggaagcagtc
600

caatgagaaa tggcaagaca tcatcaagga ggtgcgggttc ttacagaagc tccggcatcc
660

10 caacaccatt cagtaccggg gctgttacct gagggagcac acggcttggc tggtaatgga
720

gtattgcctg ggctcagctt ctgaccttct agaagtgcac aagaaacccc ttcaggaggt
780

15 agagatcgca gctgtgaccc acggggcgct tcagggcctg gcatatctgc actcccacaa
840

catgatccat agggatgtga aggctggaaa catcctgctg tcagagccag ggtagtgaa
900

gctaggggac tttggttctg cgtccatcat ggcacctgcc aactccttcg tgggcacccc
960

25 atactggatg gcacccgagg tgatcctggc catggatgag gggcagtacg atggcaaagt
1020

ggacgtctgg tccttgggga taacctgcat cgagctggct gaacggaaac caccgctctt
1080

30 taacatgaat gcgatgagt ccttatacca cattgcacag aacgaatccc ccgtgctcca
1140

gtcaggacac tggcttgagt acttccggaa tttgtcgac tcctgtcttc agaaaatccc
1200

tcaagacaga ccaacctcag aggttctcct gaagcaccgc tttgtgctcc gggagcggcc
1260

40 acccacagtc atcatggacc tgatccagag gaccaaggat gccgtgcggg agctggacaa
1320

cctgcagtac cgcaagatga agaagatcct gttccaagag gcacccaacg gccctggtgc
1380

45 cgaggcccca gaggaggaag aggaggccga gccctacatg caccggggccg ggactctgac
1440

cagcctcgag agtagccact cagtgccag catgtccatc agcgcctcca gccagagcag
1500

ctccgtcaac agcctagcag atgcctcaga caacgaggaa gaggaggagg aggaggagga
1560

55 agaggaggag gaggaagaag gccctgaagc ccgggagatg gccatgatgc aggaggggga
1620

gcacacagtc acctctcaca gctccattat ccaccggctg ccgggctctg acaacctata
1680

5 tgatgacccc taccagccag agataacccc cagccctctc cagccgcctg cagccccagc
1740

tcccacttcc accacctctt ctgcccgcgc ccgggcctac tgccgtaacc gagaccatt
1800

10 tgccaccatc cgaaccgcct ccctggtcag ccgtcagatc caggagcatg agcaggactc
1860

tgcgctgcgg gagcagctga gcggctataa gcggatgcga cgacagcacc agaagcagct
1920

15 gctggcccctg gagtcacggc tgaggggtga acgggaggag cacagtgcac ggctgcagcg
1980

ggagcttgag gcgcagcggg ctggctttgg ggagaggca gaaaagctgg cccggcggca
2040

20 ccaggccata ggtgagaagg aggcacgagc tgcccaggcc gaggagcgga agttccagca
2100

25 gcacatcctt gggcagcaga agaaggagct ggctgccctg ctggaggcac agaagcggac
2160

ctacaaactt cgcaaggaac agctgaagga ggagctccag gagaacccca gcactcccaa
2220

30 gcgggagaag gccgagtggc tgctgcggca gaaggagcag ctccagcagt gccaggcgga
2280

ggaggaagca gggctgctgc ggcggcagcg ccagtacttt gagctgcagt gtcgccagta
2340

35 caagcgcaag atgttgctgg ctccggcacag cctggaccag gacctgctgc gggaggacct
2400

40 gaacaagaag cagaccaga aggacttga gtgtgactg ctgcttcggc agcacgaggc
2460

cacgcgggag ctggagctgc ggcagctcca ggccgtgcag cgcacgcggg ctgagctcac
2520

45 ccgcctgcag caccagacgg agctgggcaa ccagctggag tacaacaagc ggcgtgagca
2580

agagttgcgg cagaagcatg cggcccaggt tcgccagcag cccaagagcc tcaaacttaa
2640

50 ggagctgcag atcaagaagc agttccagga gacgtgtaag atccagactc ggcagtacaa
2700

55 ggctctgcga gcacacttgc tggagaccac gcccaaagct cagcacaaga gcctccttaa
2760

gcgggtcaag gaagagcaga cccgcaagct ggcgatcttg gcggagcagt atgaccagtc
2820

5 catctcagag atgctcagct cacaggcgct gcggcttgat gagaccagg aggcagagtt
2880
ccaggccctt cggcagcagc ttcaacagga gctggagctg ctcaacgctt accagagcaa
2940

10 gatcaagatc cgcacagaga gccagcacga gagggagctg cgggagctgg agcagagggt
3000
cgcgctgcgg cgggcactgc tggagcagcg ggtggaagag gagctgctgg ccctgcagac
3060

15 aggacgtcc gagcgaatcc gcagtctgct tgagcggcag gccctgaga tcgaggcctt
3120
cgatgcggaa agcatgaggc tgggcttctc cagcatggct ctggggggca tcccggctga
3180
agctgctgcc cagggtatc ctgctccacc ccctgcccc gcttggccct cccgtcccgt
3240

25 tccccgttct ggggcacact ggagccatgg ccctcctoca ccaggcatgc cccctccagc
3300
ctggcgctcag ccgtctctgc tggctcccc agggccccca aactggctgg ggccccccac
3360

30 acaaagtggg acaccccgctg gcggagccct gctgctgcta agaaacagcc cccagcccct
3420
gcggcgggca gcctcggggg gcagtggcag tgagaatgtg ggccccctg ctgccgcggt
3480
gccccggccc ctgagccgca gcaccagtgt cgcttccac atcctcaatg gttcttccca
3540

40 cttctattcc tgagggtcag cggggaggag cagatgagct gggcagggca ggggtgggtg
3600
gagcctgacc ctggagggca ctgagctgga ggcccctgca agggtagggg acaagatgta
3660

45 ggctccagct cccctcagac ctctcatct catgagcttc ttggggctgg ccagtggccc
3720
agggccagct tggcgataga tgcctcaagg ctgcctggga gccccgctc cctaccatgg
3780
tgccaggggt ctccctccgc cacctaggaa aggagggaga tgtgcgtgtc aaatattcat
3840

55 ctagtcccct gggggagggg aagggtgggt ctagacatac tatattcaga gaactatact
3900

accctcacag tgaggccctc agacctgcca cagggcagag caggtctggg gcctgaggca
3960

gggagaaatga gaggccacct tactggcagg aaggatcagg atggggctctt ggggtcagga
5 4020

tgcttgggtc tcttcccgtc actgtctgac gtctgtgcc gtcttgcct ttatcttttt
4080

10 tttttttttt taattgggat cagggctggg gcggggaaac aaggaagga ccttgaagg
4140

ggctgtccc aggcctgggg ggcagtcgtg ggagcccctc tcagctgtgg ggctggcaca
4200

15 gagccccagg caagctttta ataaactgtt ggttattcta ac
4242

20 <210> .8

<211> 1049

<212> PRT

25 <213> Homo sapiens

30 <300>

<308> Genbank NP_004774.1

<309> 2000-11-01

35 <313> (1)..(1049)

40 <400> 8

Met Pro Ala Gly Gly Arg Ala Gly Ser Leu Lys Asp Pro Asp Val Ala
1 5 10 15

45 Glu Leu Phe Phe Lys Asp Asp Pro Glu Lys Leu Phe Ser Asp Leu Arg
20 25 30

50 Glu Ile Gly His Gly Ser Phe Gly Ala Val Tyr Phe Ala Arg Asp Val
35 40 45

55 Arg Asn Ser Glu Val Val Ala Ile Lys Lys Met Ser Tyr Ser Gly Lys
50 55 60

Gln Ser Asn Glu Lys Trp Gln Asp Ile Ile Lys Glu Val Arg Phe Leu
 65 70 75 80

5 Gln Lys Leu Arg His Pro Asn Thr Ile Gln Tyr Arg Gly Cys Tyr Leu
 85 90 95

10 Arg Glu His Thr Ala Trp Leu Val Met Glu Tyr Cys Leu Gly Ser Ala
 100 105 110

15 Ser Asp Leu Leu Glu Val His Lys Lys Pro Leu Gln Glu Val Glu Ile
 115 120 125

Ala Ala Val Thr His Gly Ala Leu Gln Gly Leu Ala Tyr Leu His Ser
 130 135 140

20 His Asn Met Ile His Arg Asp Val Lys Ala Gly Asn Ile Leu Leu Ser
 145 150 155 160

25 Glu Pro Gly Leu Val Lys Leu Gly Asp Phe Gly Ser Ala Ser Ile Met
 165 170 175

30 Ala Pro Ala Asn Ser Phe Val Gly Thr Pro Tyr Trp Met Ala Pro Glu
 180 185 190

35 Val Ile Leu Ala Met Asp Glu Gly Gln Tyr Asp Gly Lys Val Asp Val
 195 200 205

Trp Ser Leu Gly Ile Thr Cys Ile Glu Leu Ala Glu Arg Lys Pro Pro
 210 215 220

40 Leu Phe Asn Met Asn Ala Met Ser Ala Leu Tyr His Ile Ala Gln Asn
 225 230 235 240

45 Glu Ser Pro Val Leu Gln Ser Gly His Trp Ser Glu Tyr Phe Arg Asn
 245 250 255

50 Phe Val Asp Ser Cys Leu Gln Lys Ile Pro Gln Asp Arg Pro Thr Ser
 260 265 270

55 Glu Val Leu Leu Lys His Arg Phe Val Leu Arg Glu Arg Pro Pro Thr
 275 280 285

Val Ile Met Asp Leu Ile Gln Arg Thr Lys Asp Ala Val Arg Glu Leu

290

295

300

5 Asp Asn Leu Gln Tyr Arg Lys Met Lys Lys Ile Leu Phe. Gln Glu Ala
 305 310 315 320

10 Pro Asn Gly Pro Gly Ala Glu Ala Pro Glu Glu Glu Glu Glu Ala Glu
 325 330 335

15 Pro Tyr Met His Arg Ala Gly Thr Leu Thr Ser Leu Glu Ser Ser His
 340 345 350

20 Ser Val Pro Ser Met Ser Ile Ser Ala Ser Ser Gln Ser Ser Ser Val
 355 360 365

25 Asn Ser Leu Ala Asp Ala Ser Asp Asn Glu Glu Glu Glu Glu Glu
 370 375 380

30 Glu Glu Glu Glu Glu Glu Glu Gly Pro Glu Ala Arg Glu Met Ala
 385 390 395 400

35 Met Met Gln Glu Gly Glu His Thr Val Thr Ser His Ser Ser Ile Ile
 405 410 415

40 His Arg Leu Pro Gly Ser Asp Asn Leu Tyr Asp Asp Pro Tyr Gln Pro
 420 425 430

45 Glu Ile Thr Pro Ser Pro Leu Gln Pro Pro Ala Ala Pro Ala Pro Thr
 435 440 445

50 Ser Thr Thr Ser Ser Ala Arg Arg Arg Ala Tyr Cys Arg Asn Arg Asp
 450 455 460

55 His Phe Ala Thr Ile Arg Thr Ala Ser Leu Val Ser Arg Gln Ile Gln
 465 470 475 480

60 Glu His Glu Gln Asp Ser Ala Leu Arg Glu Gln Leu Ser Gly Tyr Lys
 485 490 495

65 Arg Met Arg Arg Gln His Gln Lys Gln Leu Leu Ala Leu Glu Ser Arg
 500 505 510

70 Leu Arg Gly Glu Arg Glu Glu His Ser Ala Arg Leu Gln Arg Glu Leu
 515 520 525

Gln Phe Gln Glu Thr Cys Lys Ile Gln Thr Arg Gln Tyr Lys Ala Leu
 755 760 765

5 Arg Ala His Leu Leu Glu Thr Thr Pro Lys Ala Gln His Lys Ser Leu
 770 775 780

10 Leu Lys Arg Leu Lys Glu Glu Gln Thr Arg Lys Leu Ala Ile Leu Ala
 785 790 795 800

15 Glu Gln Tyr Asp Gln Ser Ile Ser Glu Met Leu Ser Ser Gln Ala Leu
 805 810 815

20 Arg Leu Asp Glu Thr Gln Glu Ala Glu Phe Gln Ala Leu Arg Gln Gln
 820 825 830

25 Leu Gln Gln Glu Leu Glu Leu Leu Asn Ala Tyr Gln Ser Lys Ile Lys
 835 840 845

30 Ile Arg Thr Glu Ser Gln His Glu Arg Glu Leu Arg Glu Leu Glu Gln
 850 855 860

35 Arg Val Ala Leu Arg Arg Ala Leu Leu Glu Gln Arg Val Glu Glu Glu
 865 870 875 880

40 Leu Leu Ala Leu Gln Thr Gly Arg Ser Glu Arg Ile Arg Ser Leu Leu
 885 890 895

45 Glu Arg Gln Ala Arg Glu Ile Glu Ala Phe Asp Ala Glu Ser Met Arg
 900 905 910

50 Leu Gly Phe Ser Ser Met Ala Leu Gly Gly Ile Pro Ala Glu Ala Ala
 915 920 925

55 Ala Gln Gly Tyr Pro Ala Pro Pro Pro Ala Pro Ala Trp Pro Ser Arg
 930 935 940

60 Pro Val Pro Arg Ser Gly Ala His Trp Ser His Gly Pro Pro Pro Pro
 945 950 955 960

65 Gly Met Pro Pro Pro Ala Trp Arg Gln Pro Ser Leu Leu Ala Pro Pro
 965 970 975

Gly Pro Pro Asn Trp Leu Gly Pro Pro Thr Gln Ser Gly Thr Pro Arg
 980 985 990

5 Gly Gly Ala Leu Leu Leu Leu Arg Asn Ser Pro Gln Pro Leu Arg Arg
 995 1000 1005

10 Ala Ala Ser Gly Gly Ser Gly Ser Glu Asn Val Gly Pro Pro Ala
 1010 1015 1020

15 Ala Ala Val Pro Gly Pro Leu Ser Arg Ser Thr Ser Val Ala Ser
 1025 1030 1035

20 His Ile Leu Asn Gly Ser Ser His Phe Tyr Ser
 1040 1045

25 <210> 9
 <211> 3285
 <212> DNA
 <213> Homo sapiens

30 <220>
 <221> misc_feature

35 <222> (3245)..(3245)
 <223> a, t, c or g

40 <220>
 <221> misc_feature

45 <222> (3263)..(3263)
 <223> a, t, c or g

50 <220>
 <221> misc_feature

55 <222> (3278)..(3278)
 <223> a, t, c or g

<220>
5 <221> misc_feature
<222> (72)..(72)
10 <223> a, t, c or g

<300>
15 <308> EMBL AY049015
<309> 2001-10-15
20 <313> (1)..(3285)

<400> 9
25 caacggattt catttcatac agatgaacca aggatcggga tagcagtata aaattagaat
60
caagacagct gncgtccaag caggatgcc tcaactaaca gagcaggcag cctgaaggac
120
30 cctgaaattg cagagctctt cttcaaagaa gatccagaga agctcttcac agatctcaga
180
gaaattggcc atggaagctt tggagcagtg tattttgcac gagatgtgcg taccaatgaa
35 240
gtggtggcca tcaagaaaat gtcttatagt ggaaagcagt ctactgagaa atggcaggat
300
40 attattaagg aagtcaagtt tctacaaaga ataaaacatc ccaacagtat agaatacaaa
360
ggctgttatt tacgtgaaca cacagcatgg cttgtaatgg aatattgttt aggatctgct
420
45 tcggatttac tagaagtica caaaaagcca ttacaagaag tggaaatagc agcaattaca
480
catggtgctc ttcagggatt agcctactta cattctcata ctatgattca tagagatatc
50 540
aaagcaggaa atatccttct gacagaacca ggccaggatga aacttgctga ctttggctct
600
55 gcttcacatg catcacctgc caattccttt gtgggaacgc cgtattggat ggccccagaa
660

gtaatttttag ccatggatga aggacaatat gatggcaaag tagatgtgtg gtctcttggg
720

5 ataacatgta ttgaactagc ggaaaggaag cctcctttat ttaatatgaa tgcaatgagt
780

gccttatatc acatagccca aatgaatcc cctacactac agtctaata ga atggctctgat
840

10 tatttttcgca actttgtaga ttcttgccctc cagaaaatcc ctcaagatcg acctacatca
900

gaggaacttt taaagcacat atttggttctt cgggagcgcc ctgaaaccgt gttaatagat
960

15 ctcatcaga ggacaaagga tgcagtaaga gagctggaca atctgcagta tcgaaagatg
1020

20 aagaaactcc ttttccagga ggcacataat ggaccagcag tagaagcaca ggaagaagaa
1080

gaggaacaag atcatggtgt tggccggaca ggaacagtta atagtgttgg aagtaatcaa
1140

25 tccattccca gcatgtccat cagtgccagc agccaaagca gtagtgtaa cagtcttcca
1200

gatgtctcag atgacaagag tgagctagac atgatggagg gagaccacac agtgatgtct
1260

30 aacagttctg ttatccattt aaaaccagag gaagaaaatt acagagaaga gggagatcct
1320

35 agaacaagag catcagatcc acaatctcca cccaagtat ctcgtcacaa atcacactat
1380

cgtaatcgag aacactttgc tactatacgg acagcatcac tggttacgag gcaaatgcaa
1440

40 gaacatgagc aggactctga gcttagagaa caaatgtctg gctataagcg aatgaggcga
1500

caacatcaaa agcaactgat gactctggaa aacaagctaa aggctgagat ggatgaacat
1560

45 cgcctcagat tagacaaaga tcttgaaact cagcgtaaca attttgctgc agaaatggag
1620

50 aaacttatca agaaacacca ggctgctatg gagaaagagg cttaaagtga gtccaatgaa
1680

gagaaaaaat ttcagcaaca tattcaggcc caacagaaga aagaactgaa tagttttctc
1740

55 gagtcccaga aaagagagta taaacttcga aaagagcagc tttaaagagga gctaaatgaa
1800

aaccagagta cccccaaaaa agaaaaacag gaggggcttt caaagcagaa ggagaatata
 1860

5 cagcatttcc aagcagaaga agaagctaac ctctctcgac gtcaaagaca atacctagag
 1920

ctggaatgcc gtcgcttcaa gagaagaatg ttacttgggc gtcataactt agagcaggac
 1980

10 cttgtcaggg aggagttaaa caaaagacag actcagaagg acttagagca tgccatgcta
 2040

ctccgacagc atgaatctat gcaagaactg gagggtccgc acctcaacac aattcagaag
 2100

15 atgctgtgtg agttgatcag attacagcat caaactgagc tcactaacca gctggaatat
 2160

aataagcgaa gagaacgaga actaagacga aagcatgtca tggaagttcg acaacagcct
 2220

20 aagagtttga agtctaaaga actccaaata aaaaagcagt ttcaggatac ctgcaaaatc
 2280

25 caaaccagac agtacaaagc attaagaaat cacctgctgg agactacacc aaagagtgag
 2340

cacaaagctg ttctgaaacg gctcaaggag gaacagaccc ggaaattagc tatcttggct
 2400

30 gagcagtatg atcacagcat taatgaaatg ctctccacac aagccctgcg ttggatgaa
 2460

gcacaggaag cagagtgcc a ggttttgaag atgcagctgc agcaggaact ggagctgttg
 2520

aatgcgtatc agagcaaaat caagatgcaa gctgaggcac aacatgatcg agagcttcgc
 2580

40 gagcttgaac agaggggtctc cctccggagg gcaactcttag aacaaaagat tgaagaagag
 2640

atgttggctt tgcagaatga gcgcaçagaa cgaatacgaa gcctgttigga acgtcaagcc
 2700

45 agagagattg aagcttttga ctctgaaagc atgagactag gttttagtaa tatggtgctt
 2760

tctaattctt cccctgaggc attcagccac agctaccgag gagcttctgg ttggtcacac
 2820

50 aaccctactg ggggtccagg acctcactgg ggtcatccca tgggtggccc accacaagct
 2880

55 tggggccatc caatgcaagg tggacccag ccattgggtc accttcagg gccaatgcaa
 2940

ggggtacctc gaggtagcag tatgggagtc cgcaatagcc cccaagctct gaggcggaca
3000

5 gcttctgggg gacggacgga gcagggcatg agcagaagca cgagtgtcac ttcacaaata
3060

tccaatgggt cacacatgtc ttatacataa cttataatt gagagtggca attccgctgg
3120

10 agctgtctgc caaaagaaac tgcctacaga catcatcaca gcagcctcct cacttgggta
3180

ctacagtgtg gaagctgagt gcatatggta tattttattc atttttgtaa agcgttctgt
3240

15 ttgngttta ctaattggga tgncatagta cttggctncc cgggt
3285

20 <210> 10

<211> 1005

<212> PRT

25 <213> Homo sapiens

30 <300>

<308> SpTrEMBL

<309> 2000-10-01

35 <313> (1)..(1005)

40 <400> 10

Leu Leu Ser Arg Met Pro Ser Thr Asn Arg Ala Gly Ser Leu Lys Asp
1 5 10 15

45 Pro Glu Ile Ala Glu Leu Phe Phe Lys Glu Asp Pro Glu Lys Leu Phe
20 25 30

50 Thr Asp Leu Arg Glu Ile Gly His Gly Ser Phe Gly Ala Val Tyr Phe
35 40 45

Ala Arg Asp Val Arg Thr Asn Glu Val Val Ala Ile Lys Lys Met Ser
50 55 60

Tyr Ser Gly Lys Gln Ser Thr Glu Lys Trp Gln Asp Ile Ile Lys Glu
65 70 75 80

5 Val Lys Phe Leu Gln Arg Ile Lys His Pro Asn Ser Ile Glu Tyr Lys
85 90 95

10 Gly Cys Tyr Leu Arg Glu His Thr Ala Trp Leu Val Met Glu Tyr Cys
100 105 110

15 Leu Gly Ser Ala Ser Asp Leu Leu Glu Val His Lys Lys Pro Leu Gln
115 120 125

Glu Val Glu Ile Ala Ala Ile Thr His Gly Ala Leu Gln Gly Leu Ala
130 135 140

20 Tyr Leu His Ser His Thr Met Ile His Arg Asp Ile Lys Ala Gly Asn
145 150 155 160

25 Ile Leu Leu Thr Glu Pro Gly Gln Val Lys Leu Ala Asp Phe Gly Ser
165 170 175

30 Ala Ser Met Ala Ser Pro Ala Asn Ser Phe Val Gly Thr Pro Tyr Trp
180 185 190

35 Met Ala Pro Glu Val Ile Leu Ala Met Asp Glu Gly Gln Tyr Asp Gly
195 200 205

Lys Val Asp Val Trp Ser Leu Gly Ile Thr Cys Ile Glu Leu Ala Glu
210 215 220

40 Arg Lys Pro Pro Leu Phe Asn Met Asn Ala Met Ser Ala Leu Tyr His
225 230 235 240

45 Ile Ala Gln Asn Glu Ser Pro Thr Leu Gln Ser Asn Glu Trp Ser Asp
245 250 255

50 Tyr Phe Arg Asn Phe Val Asp Ser Cys Leu Gln Lys Ile Pro Gln Asp
260 265 270

Arg Pro Thr Ser Glu Glu Leu Leu Lys His Ile Phe Val Leu Arg Glu
275 280 285

55

Arg Pro Glu Thr Val Leu Ile Asp Leu Ile Gln Arg Thr Lys Asp Ala

290

295

300

5 Val Arg Glu Leu Asp Asn Leu Gln Tyr Arg Lys Met Lys Lys Leu Leu
305 310 315 320

10 Phe Gln Glu Ala His Asn Gly Pro Ala Val Glu Ala Gln Glu Glu Glu
325 330 335

Glu Glu Gln Asp His Gly Val Gly Arg Thr Gly Thr Val Asn Ser Val
340 345 350

15 Gly Ser Asn Gln Ser Ile Pro Ser Met Ser Ile Ser Ala Ser Ser Gln
355 360 365

20 Ser Ser Ser Val Asn Ser Leu Pro Asp Val Ser Asp Asp Lys Ser Glu
370 375 380

25 Leu Asp Met Met Glu Gly Asp His Thr Val Met Ser Asn Ser Ser Val
385 390 395 400

Ile His Leu Lys Pro Glu Glu Glu Asn Tyr Arg Glu Glu Gly Asp Pro
405 410 415

30 Arg Thr Arg Ala Ser Asp Pro Gln Ser Pro Pro Gln Val Ser Arg His
420 425 430

35 Lys Ser His Tyr Arg Asn Arg Glu His Phe Ala Thr Ile Arg Thr Ala
435 440 445

40 Ser Leu Val Thr Arg Gln Met Gln Glu His Glu Gln Asp Ser Glu Leu
450 455 460

45 Arg Glu Gln Met Ser Gly Tyr Lys Arg Met Arg Arg Gln His Gln Lys
465 470 475 480

Gln Leu Met Thr Leu Glu Asn Lys Leu Lys Ala Glu Met Asp Glu His
485 490 495

50 Arg Leu Arg Leu Asp Lys Asp Leu Glu Thr Gln Arg Asn Asn Phe Ala
500 505 510

55 Ala Glu Met Glu Lys Leu Ile Lys Lys His Gln Ala Ala Met Glu Lys
515 520 525

Glu Ala Lys Val Met Ser Asn Glu Glu Lys Lys Phe Gln Gln His Ile
530 535 540

Gln Ala Gln Gln Lys Lys Glu Leu Asn Ser Phe Leu Glu Ser Gln Lys
545 550 555 560

Arg Glu Tyr Lys Leu Arg Lys Glu Gln Leu Lys Glu Glu Leu Asn Glu
565 570 575

Asn Gln Ser Thr Pro Lys Lys Glu Lys Gln Glu Trp Leu Ser Lys Gln
580 585 590

Lys Glu Asn Ile Gln His Phe Gln Ala Glu Glu Glu Ala Asn Leu Leu
595 600 605

Arg Arg Gln Arg Gln Tyr Leu Glu Leu Glu Cys Arg Arg Phe Lys Arg
610 615 620

Arg Met Leu Leu Gly Arg His Asn Leu Glu Gln Asp Leu Val Arg Glu
625 630 635 640

Glu Leu Asn Lys Arg Gln Thr Gln Lys Asp Leu Glu His Ala Met Leu
645 650 655

Leu Arg Gln His Glu Ser Met Gln Glu Leu Glu Phe Arg His Leu Asn
660 665 670

Thr Ile Gln Lys Met Arg Cys Glu Leu Ile Arg Leu Gln His Gln Thr
675 680 685

Glu Leu Thr Asn Gln Leu Glu Tyr Asn Lys Arg Arg Glu Arg Glu Leu
690 695 700

Arg	Arg	Lys	His	Val	Met	Glu	Val	Arg	Gln	Gln	Pro	Lys	Ser	Leu	Lys
705					710					715					720

Ser Lys Glu Leu Gln Ile Lys Lys Gln Phe Gln Asp Thr Cys Lys Ile
725 730 735

Gln Thr Arg Gln Tyr Lys Ala Leu Arg Asn His Leu Leu Glu Thr Thr
740 745 750

Pro Lys Ser Glu His Lys Ala Val Leu Lys Arg Leu Lys Glu Glu Gln
 755 760 765

5 Thr Arg Lys Leu Ala Ile Leu Ala Glu Gln Tyr Asp His Ser Ile Asn
 770 775 780

10 Glu Met Leu Ser Thr Gln Ala Leu Arg Leu Asp Glu Ala Gln Glu Ala
 785 790 795 800

15 Glu Cys Gln Val Leu Lys Met Gln Leu Gln Gln Glu Leu Glu Leu Leu
 805 810 815

20 Asn Ala Tyr Gln Ser Lys Ile Lys Met Gln Ala Glu Ala Gln His Asp
 820 825 830

Arg Glu Leu Arg Glu Leu Glu Gln Arg Val Ser Leu Arg Arg Ala Leu
 835 840 845

25 Leu Glu Gln Lys Ile Glu Glu Glu Met Leu Ala Leu Gln Asn Glu Arg
 850 855 860

30 Thr Glu Arg Ile Arg Ser Leu Leu Glu Arg Gln Ala Arg Glu Ile Glu
 865 870 875 880

35 Ala Phe Asp Ser Glu Ser Met Arg Leu Gly Phe Ser Asn Met Val Leu
 885 890 895

40 Ser Asn Leu Ser Pro Glu Ala Phe Ser His Ser Tyr Pro Gly Ala Ser
 900 905 910

Gly Trp Ser His Asn Pro Thr Gly Gly Pro Gly Pro His Trp Gly His
 915 920 925

45 Pro Met Gly Gly Pro Pro Gln Ala Trp Gly His Pro Met Gln Gly Gly
 930 935 940

50 Pro Gln Pro Trp Gly His Pro Ser Gly Pro Met Gln Gly Val Pro Arg
 945 950 955 960

55 Gly Ser Ser Met Gly Val Arg Asn Ser Pro Gln Ala Leu Arg Arg Thr
 965 970 975

Ala Ser Gly Gly Arg Thr Glu Gln Gly Met Ser Arg Ser Thr Ser Val
 980 985 990

5 Thr Ser Gln Ile Ser Asn Gly Ser His Met Ser Tyr Thr
 995 1000 1005

<210> 11

10 <211> 1576

<212> DNA

15 <213> artificial sequence

<220>

20 <223> RNAi fragment T17E9.1a (kin-18)

<400> 11

25 cgaaaaccag cagaagagcg aatgtcagct gaagaatgct ttcgacatcc attcattcaa
 60

cggctctcgcc catcagacac aattcaggaa ctcattcaga gaacgaâââââ tatggtatta
 120

30 gagttggata attttcaata caaaaagatg agaaaactca tgtatttgga tgaaacagaa
 180

ggaaaagaag gaagtgaagg aaatggagca tctgatgatt tagattttca tggaaatgaa
 240

35 gctaattcaa ttggaagagg tagtttttaa aattcaaagt gaaaatatta atatcttgga
 300

40 ataattttta taataattgct ttaaaccctc agcttttttt tgcagactct atcccttagt
 360

tgttcgtttt ccatctattc tcgttttcag caggagattc tgcgtcatct cgaagtgcct
 420

45 ctcttacttc tttccgatca atgcagagta gtggaggagc tggctcttta gtgtccacca
 480

atacgacggg tgctatggat aatgtgcatg gtactgtact gtttttttg ttttaggaat
 540

50 ggctttatta tttcctgcaa agttcaâââââ ttccatttat tttagttttt ctctcgaaat
 600

55 tcatcgcgca acattgagaa tttttcaââââ ttttcaggat cctctggata cggtaatgga
 660

agtagttcga cgacgagctc cgcacgccgc cgtcctccaa ttccttcgca aatgctctct
720

5 tctacatcaa cgtctggtgt tggaactatg ccgagtcatt gatcagttgg agcatcgatt
780

acggcgatcg cagtcaatcc aacaccgtct ccttcagaac ctatcccaac atcacaacca
840

10 acatcgaaat cagaatcatc ttctatactc gaaactgcac acgatgatcc tttggacacg
900

tcgatacgtg ctccagtgaag agacttgcatt atgccgcatc gagcagtcac ggaacgaata
960

15 gccacgttgc aaaatcacaa attcgcgacg cttcgttccc agagaataat caatcaggaa
1020

caagaagaat atacgaaaga gaacaatatg tatgagcaaa tgagcaagta caagcatcta
1080

cgacaagcac atcacaaga gctccaacaa tttgaagaac gatgtgcatt agatagagag
1140

25 caactgcgtg tgaaaatgga tcgagaactc gaacaattga caacgacata ctcgaaagaa
1200

aagatgagag tgaggtgttc acagaataat gaactagaca aacggaaaaa agatatcgaa
1260

30 gatggggaga aaaagatgaa aaagacgaaa aatagtcaaa atcagcagca gatgaaactg
1320

tattcagcgc aacaattgaa agaatacaag tataacaagg aggcacagaa aacagtgaga
1380

attcactttt atttgatttc tgtaaagaaa ttatacgaaa tttagacttt ataaattttt
1440

40 aaatatgaaa gttctggtca ctttttcagc tgcttctcca cttttttcaa agtttattat
1500

ttagtcttga ataatttttt aaaaaatgtc ctaaaccaag aattttcagc gattacgaag
1560

45 tctgaacatg cctcga
1576

50 <210> 12

<211> 20

<212> DNA

55 <213> artificial sequence

<220>

5 <223> primer

<400> 12
cgaaaaccag cagaagagcg
20

10

<210> 13

<211> 21

15

<212> DNA

<213> artificial sequence

20

<220>

<223> primer

25

<400> 13
tcgaggcatg ttcagacttc g
21

30

CLAIMS:

5 1. Nucleic acid, in essentially isolated form, said nucleic acid encoding one of the amino acid sequences of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8 and/or SEQ ID NO: 10, or an analog, variant, allele, ortholog, part and/or fragment of one of the amino acid sequences of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8 and/or SEQ ID NO: 10.

10 2. Nucleic acid, in essentially isolated form, said nucleic acid comprising one of the nucleotide sequences of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7 and/or SEQ ID NO: 9, or a mutant, variant, allele, analog, ortholog, part and/or fragment thereof.

15 3. Genetic construct, comprising the nucleic acid of claim 1 and/or of claim 2, and optionally one or more further elements of genetic constructs known per se.

20 4. Host cell or host organism, transformed with and/or containing a nucleic acid according to claim 1 and/or claim 2 and/or a genetic construct according to claim 3.

25 5. Host cell or host organism, (capable of) expressing and/or producing one of the amino acid sequences of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8 and/or SEQ ID NO: 10 or an analog, variant, allele, analog, ortholog, part and/or fragment thereof.

30 6. Method for producing one of the amino acid sequences of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8 and/or SEQ ID NO: 10, or an analog, variant, allele, ortholog, part and/or fragment thereof, said method comprising at least the steps of:

 a) expressing of a nucleic acid according to claim 1 and/or 2 in a suitable host cell or host organism; and optionally

b) isolating the amino acid sequence thus expressed.

7. Method for producing one of the amino acid sequences of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8 and/or SEQ ID NO: 10, or an analog, variant, allele, ortholog, part and/or fragment thereof, said method comprising at least the steps of:

a) maintaining and/or cultivating a host cell or host organism according to claim 4 or 5 under conditions such that said host cell or host organism expresses or produces one of the amino acid sequences of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8 and/or SEQ ID NO: 10, or an analog, variant, allele, ortholog, part and/or fragment thereof; and optionally

b) isolating the amino acid sequence thus expressed/produced.

8. Protein or polypeptide, in essentially isolated form, said protein or polypeptide comprising one of the amino acid sequences of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8 and/or SEQ ID NO: 10, or an analog, variant, allele, ortholog, part and/or fragment thereof.

9. Method for generating a signal that is representative for the interaction of a protein or polypeptide of claim 8 with a test chemical, said method at least comprising the steps of:

a) contacting a protein or polypeptide of claim 8, or a host cell or host organism of claim 4 or 5, with said test chemical, in such a way that a signal may be generated that is representative for the interaction between said test chemical and said protein or polypeptide; and optionally

b) detecting the signal that may thus be generated.

10. Method for identifying a modulator of a protein or polypeptide of claim 8, for example from a set or library of test chemicals, said method at least comprising the steps of:

- a) contacting a protein or polypeptide of claim 8, or a host cell or host organism of claim 4 or 5, with said test chemical, in such a way that a signal may be generated that is representative for the interaction between said test chemical and said protein or polypeptide; and optionally
- 5 b) detecting the signal that may thus be generated, said signal identifying a modulator of said amino acid sequence.

11. Modulator of a protein or polypeptide according to claim 8, identified and/or developed using a nucleic acid according to claim 1 and/or 2, a host cell according to claim 4 and/or 5, a protein or polypeptide according to claim 8, and/or a method

10 according to claim 9 and/or 10.

12. Modulator according to claim 11, being an inhibitor of a protein or polypeptide of claim 8.

15

13. Pharmaceutical composition, comprising at least one modulator according to claim 11 and/or 12 and at least one pharmaceutically acceptable carrier.

14. Pharmaceutical composition according to claim 13, being a composition

20 suitable and/or intended for oral administration.

15. Use of a modulator according to claim 11 and/or 12 in the preparation of a pharmaceutical composition.

25

16. Use of a modulator according to claim 11 and/or 12 in the preparation of a pharmaceutical composition for the prevention and/or treatment of metabolic diseases.

17. Antibody against a protein or polypeptide according to claim 8.

Figure 1

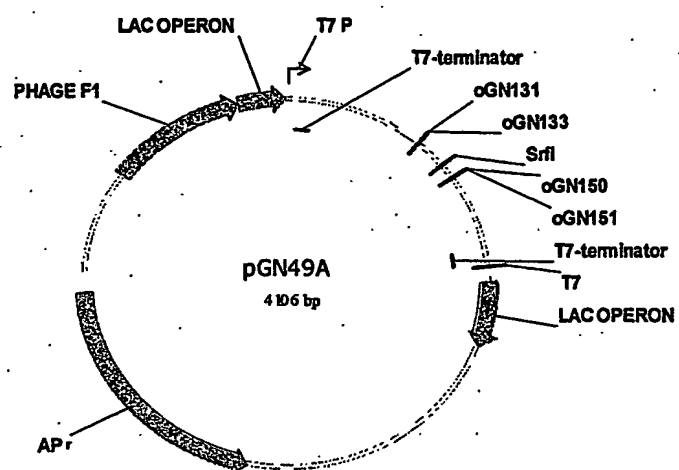


Figure 2A



Figure 2B



PCT Application
PCT/EP2003/014674



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☒ BLACK BORDERS

☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

☒ FADED TEXT OR DRAWING

☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING

☒ SKEWED/SLANTED IMAGES

☒ COLOR OR BLACK AND WHITE PHOTOGRAPHS

☐ GRAY SCALE DOCUMENTS

☐ LINES OR MARKS ON ORIGINAL DOCUMENT

☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.